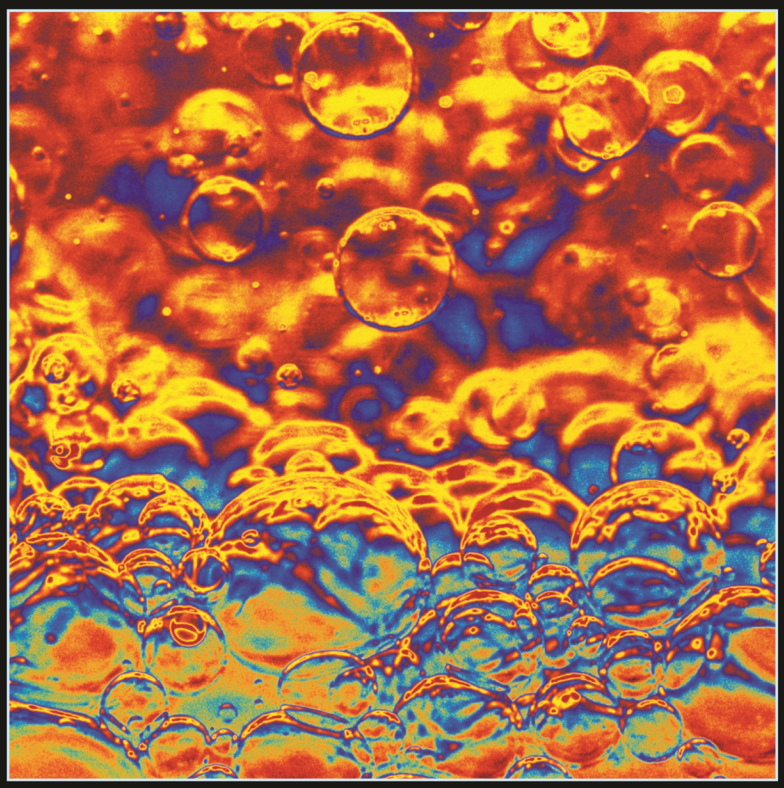
Statistical Mechanics in a Nutshell

Second Edition



Luca Peliti

Statistical Mechanics in a Nutshell

Statistical Mechanics in a Nutshell

Second Edition

Luca Peliti

PRINCETON UNIVERSITY PRESS • PRINCETON AND OXFORD

Copyright © 2024 by Princeton University Press

Princeton University Press is committed to the protection of copyright and the intellectual property our authors entrust to us. Copyright promotes the progress and integrity of knowledge. Thank you for supporting free speech and the global exchange of ideas by purchasing an authorized edition of this book. If you wish to reproduce or distribute any part of it in any form, please obtain permission.

Requests for permission to reproduce material from this work
should be sent to permissions@press.princeton.edu

Published by Princeton University Press
41 William Street, Princeton, New Jersey 08540
99 Banbury Road, Oxford OX2 6JX

press.princeton.edu

All Rights Reserved

Library of Congress Cataloging-in-Publication Data
Names: Peliti, L. (Luca), 1948– author.
Title: Statistical mechanics in a nutshell / Luca Peliti.
Description: Second edition. | Princeton : Princeton University Press, [2024] |
Series: In a nutshell | Includes bibliographical references and index.
Identifiers: LCCN 2024006037 (print) | LCCN 2024006038 (ebook) |
ISBN 9780691248424 (hardback) | ISBN 9780691248455 (ebook)
Subjects: LCSH: Statistical mechanics. | BISAC: SCIENCE / Physics /
Quantum Theory | SCIENCE / Mechanics / Statics
Classification: LCC QC174.8 .P45 2024 (print) | LCC QC174.8 (ebook) |
DDC 530.13—dc23/eng/20240223
LC record available at https://lccn.loc.gov/2024006037
LC ebook record available at https://lccn.loc.gov/2024006038

British Library Cataloging-in-Publication Data is available

Editorial: Ingrid Gnerlich and Whitney Rauenhorst
Production Editorial: Nathan Carr
Text Design: Wanda España
Jacket/Cover Design: Wanda España
Production: Jacquie Poirier
Publicity: William Pagdatoon
Copyeditor: Bhisham Bherwani

Jacket image: Alfred Pasieka / Science Photo Library

This book has been composed in Scala

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Dedicated to the memory of Giovanni Paladin.

Contents

List of Figures xiii

List of Tables xvii

Foreword xix

Preface to the Second Edition xxi

Acknowledgments xxv

1 **Introduction** 1

1.1 The Subject Matter of Statistical Mechanics 1
1.2 Statistical Postulates 3
1.3 An Example: The Ideal Gas 4
1.4 Conclusions 9

2 **Thermodynamics** 11

2.1 Thermodynamic Systems 11
2.2 Extensive Variables 13
2.3 The Central Problem of Thermodynamics 14
2.4 Entropy 15
2.5 Simple Problems 16
2.6 Heat and Work 19
2.7 Adiabatic Transformations and the Axiomatic Foundation of the Entropy Principle 24
2.8 The Fundamental Equation 26
2.9 Energy Scheme 28
2.10 Intensive Variables and Thermodynamic Potentials 30
2.11 Free Energy and Maxwell Relations 33

2.12 Gibbs Free Energy and Enthalpy 35
2.13 The Koenig-Born Diagram 37
2.14 Other Thermodynamic Potentials 38
2.15 The Measurement of the Chemical Potential 39
2.16 The Euler and Gibbs-Duhem Equations 41
2.17 Magnetic Systems 42
2.18 Equations of State 44
2.19 Stability 45
2.20 Chemical Reactions 48
2.21 Phase Coexistence 49
2.22 The Clausius-Clapeyron Equation 51
2.23 The Coexistence Curve 52
2.24 Coexistence of Several Phases 53
2.25 The Critical Point 54
2.26 Planar Interfaces 55

3 **The Fundamental Postulate** 59
3.1 Phase Space 59
3.2 Observables 61
3.3 The Fundamental Postulate: Entropy as Phase-Space Volume 62
3.4 Liouville's Theorem 64
3.5 Quantum States 68
3.6 Systems in Contact 71
3.7 Variational Principle 72
3.8 The Ideal Gas 73
3.9 The Probability Distribution 75
3.10 The Maxwell Distribution 77
3.11 The Ising Paramagnet 78
3.12 The Canonical Ensemble 81
3.13 Generalized Ensembles 85
3.14 The *p*-*T* Ensemble 89
3.15 Quantum Ensembles 92
3.16 The Grand Canonical Ensemble 93
3.17 The Gibbs Formula for the Entropy 95
3.18 Variational Derivation of the Ensembles 97
3.19 Fluctuations of Uncorrelated Particles 99

4 **Interaction-Free Systems** 102
4.1 Harmonic Oscillators 102
4.2 Photons and Phonons 106
4.3 Boson and Fermion Gases 117
4.4 Electrons in Metals 123
4.5 Relation between Pressure and Internal Energy 127

4.6 Diamagnetism 129
4.7 White Dwarfs 135
4.8 Variational Derivation of Fermi and Bose Statistics 139
4.9 Einstein Condensation 141
4.10 Adsorption 143
4.11 Internal Degrees of Freedom 146
4.12 Chemical Equilibria in Gases 154

5 **Phase Transitions** 155
5.1 Liquid-Gas Coexistence and Critical Point 155
5.2 Van der Waals Equation 157
5.3 Binary Mixtures 161
5.4 Lattice Gas 162
5.5 The Ising Model 164
5.6 Lee-Yang Theory of Phase Transitions 165
5.7 Symmetry Breaking 169
5.8 The Order Parameter 170
5.9 Peierls's Argument 172
5.10 The One-Dimensional Ising Model 176
5.11 Duality 179
5.12 Mean-Field Theory 182
5.13 Variational Principle 186
5.14 Correlation Functions 188
5.15 The Landau Theory 192
5.16 Critical Exponents 195
5.17 The Einstein Theory of Fluctuations 196
5.18 Ginzburg Criterion 199
5.19 Limit $n \to \infty$ 201
5.20 Universality and Scaling 204
5.21 Partition Function of the Two-Dimensional Ising Model 209

6 **Renormalization Group** 215
6.1 Block Transformation 215
6.2 Decimation in the One-Dimensional Ising Model 218
6.3 Two-Dimensional Ising Model 220
6.4 Relevant and Irrelevant Operators 225
6.5 Finite Lattice Method 229
6.6 Renormalization in Fourier Space 231
6.7 Quadratic Anisotropy and Crossover 246
6.8 Critical Crossover 247
6.9 Cubic Anisotropy 252
6.10 Lower and Upper Critical Dimensions 253
6.11 Kosterlitz-Thouless Transition 255

7 **Classical Fluids** 262
- 7.1 The Partition Function of a Classical Fluid 262
- 7.2 Reduced Densities 265
- 7.3 The Density Functional Approach 276
- 7.4 Virial Expansion 281
- 7.5 Perturbation Theory 293
- 7.6 Liquid Solutions 295
- 7.7 Colloidal Suspensions 301

8 **Numerical Simulation** 305
- 8.1 Introduction 305
- 8.2 Molecular Dynamics 307
- 8.3 Thermostats in Molecular Dynamics 314
- 8.4 Monte Carlo Method 319
- 8.5 Umbrella Sampling 331
- 8.6 Accelerated Monte Carlo Methods 333
- 8.7 Discussion 336

9 **Dynamics** 339
- 9.1 Brownian Motion 339
- 9.2 Diffusion Coefficient and the Einstein Relation 342
- 9.3 Fractal Properties of Brownian Trajectories 345
- 9.4 Smoluchowski Equation 347
- 9.5 Diffusion Processes and the Fokker-Planck Equation 352
- 9.6 Correlation Functions 354
- 9.7 Kubo Formula and Sum Rules 357
- 9.8 Metastable States and Stochastic Resonance 358
- 9.9 Chemical Kinetics 362
- 9.10 Generalized Brownian Motion 365
- 9.11 Response Functions 368
- 9.12 Fluctuation-Dissipation Theorem 372
- 9.13 Onsager Reciprocity Relations 376
- 9.14 Affinities and Fluxes 378
- 9.15 Nonequilibrium Steady States 380
- 9.16 An Application 382

10 **Stochastic Thermodynamics** 386
- 10.1 Fundamentals 386
- 10.2 Thermodynamic Consistency and Stochastic Energetics 389
- 10.3 Stochastic Entropy 390
- 10.4 Simple Examples 393
- 10.5 Average Entropy Production Rate 395

10.6 Time Reversal 396
10.7 Fluctuation Relations 397
10.8 Adiabatic and Nonadiabatic Entropy Production 402
10.9 Thermodynamics of Information 404
10.10 Information Reservoirs 407
10.11 Copying Information 410
10.12 Uncertainty Relations 412

11 **Complex Systems** 416
11.1 Introduction 416
11.2 Linear Polymers in Solution 417
11.3 Percolation 425
11.4 Systems with Impurities 442
11.5 Spin Glasses 445
11.6 The Hopfield Model of Associative Memory 462

A **Probability Refresher** 468
A.1 Events and Probability 468
A.2 Random Variables 469
A.3 Averages and Moments 470
A.4 Conditional Probability: Independence 471
A.5 Characteristic Functions 473
A.6 Central Limit Theorem 474
A.7 Correlations 475
A.8 Dirac's Delta "Function" 477

B **Convex Functions and the Legendre Transformation** 479
B.1 Convex Functions 479
B.2 The Jensen Inequality 481
B.3 Legendre Transformation 482
B.4 Properties of the Legendre Transformation 484
B.5 Lagrange Multipliers 484

C **Partial and Functional Derivatives** 487
C.1 Partial Derivatives 487
C.2 Functional Derivatives 489

D **Fourier Transforms** 492
D.1 Fourier Series 492
D.2 Fourier Transform 494
D.3 Functions Defined on a Lattice: Brillouin Zones 496
D.4 Integral Representation of the Delta Function 498

E **Saddle-Point Integration** 499
E.1 Euler Integrals and the Saddle-Point Method 499
E.2 The Euler Gamma Function 501
E.3 Properties of the *N*-Dimensional Space 502

F **Basics of Information Theory** 504
F.1 Shannon Entropy 504
F.2 Mutual Information 506

G **Markov Chains** 507
G.1 Definitions 507
G.2 Spectral Properties 508
G.3 Convergence to Stationarity 509
G.4 Non-ergodic Matrices 512

H **Fundamental Physical Constants** 513

Bibliography 515

Author Index 527

Subject Index 531

Figures

1.1	Derivation of the relation between average speed and pressure	5
1.2	The Marcus and McFee experiment	8
1.3	Particle flux in the Marcus and McFee experiment	9
2.1	Systems in contact	13
2.2	Generic diagram of a thermodynamic experiment	15
2.3	A schematic representation of the fundamental surface	29
2.4	The Koenig-Born diagram for the variables entropy S and volume V	37
2.5	Koenig-Born diagram and Maxwell relations	38
2.6	Koenig-Born diagram relative to the variables S and N	38
2.7	The measurement of the chemical potential	39
2.8	Isotherm p as a function of the volume per particle v for a simple fluid	50
2.9	Isotherms of the Helmholtz and Gibbs free energies	51
2.10	Coexistence curve for a simple fluid	53
2.11	Local density profile in a liquid vapor interface	55
2.12	Coexistence of two phases, with their interface	56
3.1	Graphical derivation of the Liouville theorem	66
3.2	Phase-space distribution of two systems in contact	71
3.3	Equation of state of the Ising paramagnet	79
3.4	Specific heat of the Ising paramagnet	79
3.5	Entropy per spin of the Ising paramagnet	80
3.6	Illustration of Curie's law	88
4.1	Specific heat per oscillator in the Einstein model	106
4.2	Bose factor	109
4.3	Planck's law for the black body radiation	109
4.4	Spectrum of the cosmic background radiation	110
4.5	Specific heat per degree of freedom in the Debye model	112
4.6	Fermi factor	121
4.7	Equations of state of the ideal quantum gases	122

4.8	Plot of C/T vs. T^2 for Ba, Sr, and Ca	126
4.9	De Haas-van Alphen effect	135
4.10	The Langmuir isotherm	145
4.11	Rotational specific heat of a biatomic molecule	151
4.12	Rotational specific heat of H_2	153
5.1	Van der Waals isotherms	156
5.2	Phase diagram of CO_2 on the plane (T, p)	157
5.3	Maxwell construction	159
5.4	Coexistence curve in the vicinity of the critical temperature	160
5.5	Discontinuous transition in a model system	167
5.6	Configuration of boundaries associated with a spin configuration	174
5.7	Dual lattices	180
5.8	Scheme of the duality transformation in the $2d$ Ising model	181
5.9	Graphical solution of the mean-field equation	184
5.10	Spontaneous magnetization per spin in the Ising model	185
5.11	Specific heat per spin in an Ising model	185
5.12	Magnetic contribution to the specific heat of nickel	186
5.13	Inverse susceptibility for Dy as a function of temperature	191
5.14	Landau free energy per spin as a function of the magnetization	193
5.15	A diagram with self-intersections	210
5.16	A diagram with a three-bond vertex	211
6.1	Block transformation	216
6.2	Decimation in the $2d$ Ising model	222
6.3	Renormalized vs. original interaction in the $2d$ Ising model	224
6.4	Scheme of the critical and elementary surfaces	227
6.5	Block-spin transformation in a triangular lattice	229
6.6	Finite lattice approximation	230
6.7	Renormalization flow for the ϕ^4 model	243
6.8	Effective exponent γ_{eff}	251
6.9	Renormalization flow for the ϕ^4 model with cubic anisotropy	253
6.10	Example of a vortex and an antivortex configuration	257
6.11	Renormalization flow for the Kosterlitz-Thouless transition	261
7.1	Scheme of a neutron scattering experiment	271
7.2	The Lennard-Jones potential and its Mayer function	282
7.3	The Joule-Thomson process	283
7.4	The reduced second virial coefficient	284
7.5	Definition of an articulation point	287
7.6	The first series and bridge diagrams	291
7.7	Expansion of $c(r)$ in powers of ρ	292
7.8	Percus-Yevick approximation	292
7.9	Equation of state for the hard-sphere gas	293
7.10	Structure factor $S(k)$ for the hard-sphere gas	293
7.11	Pair distribution of a Lennard-Jones fluid	295

8.1	Illustration of finite-size scaling	331
8.2	Solution of a travelling salesman problem by simulated annealing	335
9.1	Experimental trajectory of a Brownian particle	345
9.2	A Brownian trajectory described with increasing resolution	346
9.3	Potential for a metastable state	359
9.4	Brownian particle in a bistable potential	360
9.5	Brownian particle exhibiting stochastic resonance	361
10.1	Simulation of a two-level system	394
10.2	Stochastic work and heat in a three-state system	395
10.3	Illustrating involution	398
10.4	Phase diagram of the information-fueled engine	409
11.1	Graphical expansion in the $n = 0$ model	425
11.2	Site percolation in two dimensions	426
11.3	The neighborhood of the origin in the Bethe lattice	429
11.4	Renormalization of bond percolation in the case of a 2×2 cell	438
11.5	Migdal-Kadanoff transformation	438
11.6	Examples of fractals	440
11.7	A frustrated plaquette in a two-dimensional Ising model	447
11.8	Minimal-length barrier configurations in a square lattice with frustration	448
11.9	The Gumbel distribution	451
11.10	Scheme of a pyramidal neuron in the cortex	463
B.1	Illustration of convexity	480
B.2	Legendre transform	483
B.3	Double tangent construction	483
B.4	Parametric plot of $x\,p(x) - f(x)$ vs. $p(x)$ for a non-convex function	484
B.5	Constrained optimization in the (x, y) plane	486

Tables

5.1 Classical critical exponents 196
5.2 Critical exponents of different physical systems 205
5.3 Best theoretical estimates of the critical exponents 206
6.1 Decimation in the 1d Ising model 221
A.1 Properties of the elementary probability distributions 471

Foreword

Chatting with a group of undergraduates many years ago, I asked them to rank by relative difficulty the four core subjects they had studied: classical mechanics, electromagnetism, statistical mechanics, and quantum mechanics. Quantum mechanics was rated the most difficult and classical mechanics the easiest, with a toss-up between electromagnetism and statistical mechanics for second easiest. Many thought that electromagnetism was more difficult than statistical mechanics, especially with all those "Jackson style" problems invoking tricky boundary conditions and nasty charge distributions.

I then conducted a similarly nonscientific poll of some of the denizens of the institute for theoretical physics where I allegedly worked, and still work. The ordering was almost the exact reverse. Keep in mind that we were talking about mastery at the undergraduate level, not what it might take to become a globally acknowledged authority. Everybody thought that quantum mechanics was by far the easiest, as long as you swallowed the inexplicable—you had no choice anyway. Then quantum mechanics was basically just linear algebra. Electromagnetism came in second, also merely a linear theory. Many of the professionals thought that classical mechanics could be a real brain breaker. Think of the infamous spinning top precessing on an incline plane, perhaps vibrating up and down for good measure. But surely statistical mechanics is the most difficult conceptually, with its many subtleties and paradoxes. After all, in the entire history of physics, there were only two suicides among the top brains, Boltzmann and Ehrenfest, and both of them in statistical mechanics. (Of course, there were exacerbating reasons driving them to their tragic ends: for Ehrenfest, his personal life, and for Boltzmann, that fool Mach harassing him.)

Statistical physics is hard. Thermodynamics is hard. I don't think that I understood either subject when I was an undergraduate: the professor who taught me thermodynamics never emphasized that he was supposed to specify what he was holding fixed when he differentiated.

A multitude of textbooks on statistical physics exists, and yet this book, *Statistical Mechanics in a Nutshell*, was welcomed with much joy when it first came out, and not merely because I initiated the nutshell series at Princeton University Press. The reason

worthy of celebration is that Luca Peliti has spent essentially his entire academic life thinking about statistical and thermal physics. His deep understanding is evident throughout this book. I once wrote a foreword to a textbook on quantum field theory saying something similar, that there was never a sufficient supply of textbooks on quantum field theory. I cited Feynman admonishing students that they should never put down just one coat when painting a house. Similarly, I cannot imagine anyone (besides perhaps those greats capable of actually inventing quantum field theory) learning quantum field theory from a single textbook. Those other students who told you the contrary are either kidding you or, more likely, themselves. I have always been proud to say that during my four years of graduate study, I attended Julian Schwinger's course no less than four times.

By the same token, if you were like me, you could not possibly master statistical mechanics by reading a single book. So, come learn statistical mechanics from a master of statistical physics, Luca Peliti. Believe me, you will learn a great deal.

A. Zee, Santa Barbara, California,
May 2023

Preface to the Second Edition

I learned more from my pupils than from all others.
—Lorenzo Da Ponte

This work grew out of the notes I gathered while preparing lessons in statistical mechanics—a course I had been teaching for almost 30 years at the Federico II University in Naples (Italy). The course was both introductory and advanced. As a consequence, the first part (chapters 1 to 5) of the book contains an introduction to the problems and methods of statistical mechanics, devised for students without any prior knowledge of the subject, while the second part contains an introduction to some topics of current interest. Therefore, the first group of chapters (1 to 5) are meant to be studied sequentially, while the following chapters can be studied independently of one another, with the exceptions of chapter 10—which requires some knowledge of chapter 9—and chapter 11—which depends on chapter 6.

This arrangement has led to a drastic shift in the presentation of the subject matter. I thought it necessary to begin with a modern exposition of thermodynamics, using the geometric method introduced by J. W. Gibbs and reworked by L. Tisza and H. B. Callen. This method has significant conceptual and mnemonic advantages when compared to more traditional presentations. I have chosen to introduce the fundamental postulates of statistical mechanics on the basis of the expression of entropy in terms of the volume of the accessible phase space, due to Boltzmann. In this enterprise, I follow the point of view of L. D. Landau and S.-K. Ma. This point of view allows one to do without the ergodic hypothesis, thus avoiding a series of delicate discussions that are ultimately not very relevant from a practical standpoint. In this manner, the basic techniques of statistical mechanics can be explained in their general outlines without taking too many detours. I have drastically simplified both the terminology and the formalisms used in the definition of the fundamental ensembles, relying at every step on parallels with the formulation of thermodynamics that has been previously introduced. I have experienced that this setting allows the teacher to reach in few lectures a point where the basic techniques can be used by the students to solve problems in elementary statistical mechanics (interaction-free systems). It is then possible to apply these techniques to simple physically interesting

systems like quantum gases and paramagnets, and even to less traditional problems like the elasticity of polymers.

The exploration of systems with interaction is often performed by means of *mean-field theory*. This theory is rarely mentioned in introductory statistical mechanics textbooks, and in my opinion, the space allocated to it in advanced textbooks is rarely sufficient. In my experience, it is not only fundamental, but can also be presented and understood at a beginner's level. It allows the teacher to introduce the problems connected to phase transitions and their fluctuations following a natural sequence; the phenomenological theory concerning these issues (scaling) is the final topic of the introductory part.

The second part of this work includes topics that can be studied independently of each other. The *renormalization group theory* of phase transitions represented a true conceptual revolution for our discipline. Explaining it is a notoriously difficult task, and I therefore attempt, in chapter 6, to present it in a manner that can be productive on a variety of levels. On the one hand, I explain its basic structure by introducing the concept of a Kadanoff transformation independent of its specific realizations, showing what its existence entails for the system's critical behavior. I then introduce two of the transformation's concrete realizations: one elementary, but not systematic (decimation in the Ising model), and another that is more systematic—in the sense that it can lead to a series of successive approximations—even though it requires more complex calculations. In this case, it is important to stress the principle of the method, since its actual implementation is a more advanced topic, one that is also discussed in several other texts.

I devote one chapter to the modern theory of classical fluids, which saw remarkable progress in the 1970s and 1980s. I believe that I have found a simple way to present the principles of diagrammatic expansion of the properties of fluids, as well as the derivation of integral equations. This is a good starting point to introduce more advanced methods, such as the density functional method.

Given the importance of numerical simulations in modern statistical mechanics, I devote a chapter to this topic: I spend more time on basic problems than on the actual writing of simulation programs. This is a deliberate choice because discussing these programs would require the explicit introduction of a specific programming language, the analysis of a certain number of technical aspects related to its implementation, and so on—all items that would require too much space and distract the reader. Often, the students have already encountered some of these topics in other courses, and if this is not the case, discussing them would transform the chapter into a necessarily unsatisfactory introduction to the use of computers. On the other hand, I believe that discussions of the foundations of the Monte Carlo method, of the treatment of statistical errors, and above all of the extrapolation of the thermodynamic limit are of great relevance in the context of our discipline but are often neglected in textbooks I have been able to examine. I have also emphasized the contribution that numerical simulation provides in addressing problems in the foundations of statistical mechanics—a topic that, as far as I know, has been treated only in the text by S.-K. Ma [Ma85].

Another chapter is devoted to dynamic phenomena, especially the theory of linear response and the fluctuation-dissipation theorem and its consequences, such as the Onsager reciprocity relations. My experience teaching the course has shown me that,

starting with the theory of Brownian motion and making use of the concept of generalized Brownian motion, one can rapidly and easily derive these results. This allows me to discuss, in chapter 10, one of the most exciting recent developments of the discipline, namely, stochastic thermodynamics. I present there the basic ideas of the subject, and the derivation of fluctuation relations and thermodynamic uncertainty relations in a direct and simple way. Lastly, chapter 11 contains an introduction to the study of complex systems in statistical mechanics: it discusses the theory of linear polymers and then the theory of percolation as an example of a system with frozen disorder. It then contains a short introduction to the more general theory of disordered systems and of spin glasses and allows a glimpse into one of its interdisciplinary applications—namely, the statistical theory of associative memories. In this fashion, the student can perceive the general nature of the methods of statistical mechanics, which the more traditional mechanistic approach tends to obscure.

In order to keep the size of the work within reasonable bounds, I omit the discussion of several interesting topics. More specifically, I include only a brief introduction to the kinetic theory of gases (and I do not even mention Boltzmann's equation), I do not discuss the ergodic problem, and I generally avoid discussions about the foundations of statistical mechanics. Indeed, the debate about these foundations is still ongoing but appears to me to bear little relevance to the concrete results that can be obtained within a more pragmatic approach. I also omit any reference to quantum statistical mechanics, except for those few cases that lead to calculations that are basically classical in nature. The problem is that, except for some special cases, this subject cannot be treated with the elementary methods I use in this textbook. This allows me to discuss the problems that lie at the heart of contemporary statistical mechanics, especially the theories of phase transitions and of disordered systems, with relative ease.

This second edition appears more than ten years after the first English edition. Beyond thoroughly revising the text and correcting all the errors that I could find, or that had been spotted by readers, I have taken this opportunity to make the notations more uniform and to reorganize the exposition in several places. A few new topics have been added to several chapters, and a new chapter on stochastic thermodynamics has been introduced.

This book is dedicated to the memory of Giovanni Paladin. We are still reaping the fruit of his scientific activity, so untimely and tragically interrupted.

Acknowledgments

This book could not have been written without the help of many who helped and encouraged me all along these years. I especially thank Mauro Sellitto and Mario Nicodemi, who valiantly contributed to the completion of a first draft, as well as Stanislas Leibler, whose lessons inspired me when I was giving my course its shape. Many students in my statistical mechanics course contributed both to making certain proofs clearer and to reducing the number of errors. Although I cannot thank them all here one by one, I want to express my gratitude to all of them. I also thank Antonio Coniglio, Franco Di Liberto, Rodolfo Figari, Jean-Baptiste Fournier, Silvio Franz, Giancarlo Franzese, Giuseppe Gaeta, Marc Mézard, Fulvio Peruggi, and Mohammed Saber for their observations and suggestions, as well as Giuseppe Trautteur for his encouragement and for having suggested I publish my notes in book format. Further observations and comments on the published book were made to me by Jeferson Arenzon, Alberto Imparato, Andrea Longobardo, and a number of students of my courses. I am also grateful to Alberto Clarizia, Giancarlo D'Ambrosio, Pietro Santorelli, Maurizio Serva, and Angelo Vulpiani for discussions that have led me to reconsider some points of my exposition. I also thank Luca Buonocore, Michele Campisi, Anita Mehta, Mauro Sellitto, and especially Josh Silverman for spotting mistakes in the first English edition. I warmly thank Ramgopal Agrawal, Ada Altieri, Giuseppe Gaeta, and David Lacoste for reading the manuscript of the second edition and providing me with many useful suggestions. I am especially grateful to Tony Zee for suggesting publish the book in English and for accepting to write a foreword to its second edition. Special thanks are due to Ingrid Gnerlich and Whitney Rauenhorst of Princeton University Press for their assistance. I am also grateful to Bhisham Bherwani for his careful proofreading.

I am grateful to the Naples Unit of the Istituto Nazionale di Fisica della Materia for the continuous aid they provided for my research, until the institute was suppressed by an ill-advised government decision. I also thank the Naples section of the Istituto Nazionale di Fisica Nucleare for its generous support during my years in Naples.

And to Raya, all my gratitude for her unbreakable patience and constant encouragement.

Statistical Mechanics in a Nutshell

1 Introduction

Lies, damned lies, and statistics.

—Disraeli

1.1 The Subject Matter of Statistical Mechanics

The goal of statistical mechanics is to predict the macroscopic properties of bodies, especially their thermodynamic properties, on the basis of their microscopic structure.

The macroscopic properties of greatest interest to statistical mechanics are those relating to *thermodynamic equilibrium*. As a consequence, the concept of thermodynamic equilibrium occupies a central position in the field. It is for this reason that we will first review some elements of thermodynamics, which will allow us to make the study of statistical mechanics clearer once we begin it. The examination of *nonequilibrium states* in statistical mechanics is a fairly recent development (except in the case of gases) and is currently the focus of intense research. We will omit it in this course, even though we will deal with properties that are time-dependent (but always related to thermodynamic equilibrium) in the chapter on dynamics.

The microscopic structure of systems examined by statistical mechanics can be described by means of mechanical models: for example, gases can be represented as systems of particles that obey the classical equations of motion and interact by means of a phenomenologically determined potential. Other examples of mechanical models are those that represent polymers as a chain of interconnected particles, or the classical model of crystalline systems, in which particles are arranged in space according to a regular pattern, and oscillate around the minimum of the potential energy due to their mutual interaction. The models we use are, however, rather abstract and often exhibit only a faint resemblance to the basic mechanical description (more specifically, to the quantum nature of matter). How such abstract models are able to describe the behavior of actual systems is itself one of the more interesting questions of statistical mechanics, and has led to establishing the theory of universality and its foundation in the renormalization group.

The models of systems dealt with by statistical mechanics have some common characteristics. We are in any case dealing with systems with a large number of degrees of

freedom: the reason lies in the corpuscular (atomic) nature of matter. Avogadro's constant, $N_A \simeq 6.022 \cdot 10^{23}\ \mathrm{mol}^{-1}$—that is, the number of molecules contained in a gram-mole (mole)—provides us with an order of magnitude of the degrees of freedom contained in a thermodynamic system. (The values of this and other fundamental physical constants are given in sec. H.) The degrees of freedom that one considers should have more or less comparable effects on the global behavior of the system.

Exercise 1.1 (On Avogadro's number) Imagine we could "mark" the water molecules contained in a small flask of 100 cc, and pour them into the sea. If we fill the flask from the sea after having waited for the flask's "marked" molecules to have distributed uniformly in the oceans, how many of the molecules can we expect to find back on average?

NOTE. The surface of the oceans equals 71% of the earth's surface, and its mean depth is 3,800 m. The molecular weight of water is equal to 18.

This state of affairs excludes the application of the methods of statistical mechanics to cases in which a restricted number of degrees of freedom "dominates" the others—for example, in celestial mechanics, although the number of degrees of freedom of the planetary system is immense, an approximation in which each planet is considered as a particle is a good start. In this case, we can state that the translational degrees of freedom (three per planet)—possibly with the addition of the rotational degrees of freedom, also a finite number—dominate all others. It follows from these considerations that one encounters quite hard problems if one naively attempts to apply statistical concepts to human sciences, such as politics. Indeed, even if a nation's political system includes a very high number of degrees of freedom, it is possible to identify some degrees of freedom that are much more important than the rest. On the other hand, statistical methods can also be applied to systems that are not strictly speaking mechanical—for example, neural networks (understood as models of the brain's components), urban thoroughfares (traffic models), and some problems of a geometric nature (for example, percolation).

The simplest statistical mechanical model is that of a large number of identical particles, free of mutual interaction, inside a container with impenetrable and perfectly elastic walls. This is the model of the **ideal gas**, which describes quite well the behavior of real gases at low densities, and more specifically allows one to derive the well-known equation of state.

The introduction of pair interactions between the particles of the ideal gas allows us to obtain the standard model for **simple fluids**. Generally speaking, this model cannot be resolved exactly and is studied by means of perturbation or numerical techniques. It allows one to describe the behavior of real gases (especially rare gases), and the liquid-vapor transition (boiling and condensation).

The preceding models are of a classical (that is, not quantum) nature and can be applied only when the temperatures are not too low. The quantum effects that follow from the inability to distinguish particles from one another are very important, and can be dealt with at the introductory level if one omits interactions between particles. In this fashion, we obtain models for **quantum gases**, further distinguished as **fermions** or **bosons**, depending on the nature of the particles.

The model of **noninteracting fermions** describes the behavior of conduction electrons in metals fairly well, once one redefines the dependence of their energy on their momentum in a suitable way. The thermodynamic properties are governed by the Pauli exclusion principle.

The model of **noninteracting bosons** has two important applications: radiating energy in a cavity (also known as **black body**) can be conceived as a set of particles (**photons**) that are bosonic in nature; moreover, helium (whose most common isotope, He, is bosonic in nature) exhibits, at low temperatures, a remarkable series of properties that can be interpreted on the basis of the noninteracting boson model. Actually, the transition of ^{4}He to a superfluid state, also referred to as the λ **transition**, is connected to the **Einstein condensation**, which occurs in a gas of noninteracting bosons at high densities. Obviously, interactions between helium atoms are not negligible, but their effects can be studied by means of analytic methods such as perturbation theory.

In many of the statistical models we will describe, however, the system's fundamental elements will not be "particles," and the fundamental degrees of freedom will not be mechanical (position and velocity or impulse). If we want to understand the origin of ferromagnetism, for example, we focus on the degrees of freedom that are most relevant for the phenomenon, such as the orientation of the magnetic moments of the electrons, called their **spin**. The spin, being of a quantum nature, can assume only a finite number of values. The simplest case is when there are only two values—in this fashion, we obtain a simple model of ferromagnetism, known as the **Ising model**, which is by far the most studied model in statistical mechanics. The ferromagnetic solid is therefore represented as a regular lattice in space, where for each point of the lattice there is a spin variable that can assume the values $+1$ and -1. This model allows one to describe the paramagnet-ferromagnet transition, as well as other similar transitions.

1.2 Statistical Postulates

The behavior of a mechanical system is determined not only by its structure, represented by the equations of motion, but also by its initial conditions. Therefore the laws of mechanics are not enough by themselves to define the behavior of a mechanical system that contains a large number of degrees of freedom if nothing is said about the relevant initial conditions. It is therefore necessary to complete the description of the system with some additional postulates—the statistical postulates in the strict sense of the word—that concern these initial conditions.

The path to arrive at the formulation of statistical postulates is fairly twisted. In the following section, we discuss the relatively simple case of an ideal gas. We conjecture the distribution of the positions and velocities of particles in an ideal gas at equilibrium, following Maxwell's reasoning in a famous article [Maxw60], and we see how the equation of state of the ideal gas follows from this conjecture and the laws of mechanics. What this argument does not prove is that this distribution is conserved—in other words, if the equilibrium distribution of positions and velocities of the particles holds at a certain instant in time, it also holds at each following instant, as the system evolves according to its

equations of motion. It is also necessary to show that if the initial distribution is different from the equilibrium one, the actual distribution approaches the equilibrium one as a consequence of the system's evolution—that is, that the system "naturally" approaches equilibrium. Boltzmann's great contribution was to state clearly the problems mentioned above and to make a bold attempt at their solution.

1.3 An Example: The Ideal Gas

1.3.1 Definition of the System

In the model of an ideal gas, one considers N point-like bodies (or particles), with mass equal to m, identical to one another, free of mutual interaction, inside a container of volume V whose walls are perfectly reflecting. The mechanical state of the system is identified by the position vector $\boldsymbol{r}_i$ and the velocity vector $\boldsymbol{v}_i$ (both three-dimensional vectors) of each particle i. These vectors evolve according to the laws of mechanics.

1.3.2 Maxwell's Postulates

The assumption is that the vectors are distributed "randomly," and more specifically that:

1. The velocity and position vectors pertaining to different particles are independent of one another. This hypothesis certainly does not apply, among other examples, to particles that are very close to each other, because the position of two particles that are very close is undoubtedly influenced by the forces that act between them. One can, however, expect that if the gas is very diluted, deviations from this hypothesis will have negligible consequences. If one accepts the hypothesis of independent particles, we can describe the state of the system by giving, for each value of $\boldsymbol{r} = (x, y, z)$ and $\boldsymbol{v} = (v_x, v_y, v_z)$, the number $\mathrm{d}N$ of particles whose position is located within a box, with sides $\mathrm{d}\boldsymbol{r} = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z)$ placed around $\boldsymbol{r}$, and that are simultaneously driven by a velocity whose vector lies in a box of sides $\mathrm{d}\boldsymbol{v} = (\mathrm{d}v_x, \mathrm{d}v_y, \mathrm{d}v_z)$ around the vector $\boldsymbol{v}$: we then have $\mathrm{d}N = f(\boldsymbol{r}, \boldsymbol{v})\, \mathrm{d}\boldsymbol{r}\, \mathrm{d}\boldsymbol{v}$. This defines the single-particle probability distribution $f(\boldsymbol{r}, \boldsymbol{v})$.

2. Position is independent of velocity (in the sense given by probability theory), and therefore the probability distribution $f(\boldsymbol{r}, \boldsymbol{v})$ factorizes: $f(\boldsymbol{r}, \boldsymbol{v}) = f^{(r)}(\boldsymbol{r}) f^{(v)}(\boldsymbol{v})$.

3. Density is uniform in the space occupied by the gas, and therefore $f^{(r)}(\boldsymbol{r}) = N/V = \rho = \text{const.}$ if $\boldsymbol{r}$ is inside the container, and equal to zero otherwise.

4. The velocity components are mutually independent, and we have therefore $f^{(v)}(\boldsymbol{v}) = f^{(x)}(v_x) f^{(y)}(v_y) f^{(z)}(v_z)$.

5. The distribution $f^{(v)}(\boldsymbol{v})$ is isotropic in velocity space, so that $f^{(v)}(\boldsymbol{v})$ depends in actual fact only on the magnitude $v = |\boldsymbol{v}|$ of $\boldsymbol{v}$.

The basic properties of probability distributions are summarized in sec. A.

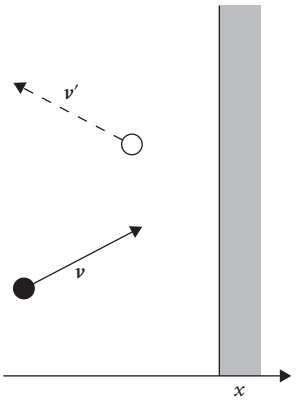


FIGURE 1.1. Derivation of the relation between average speed and pressure. The particle driven by velocity $\boldsymbol{v}$ is about to hit the wall. After the impact, it will have velocity $\boldsymbol{v}'$, in which the x component (normal to the wall) has changed sign.

Exercise 1.2 Prove that the only distribution that satisfies postulates 4 and 5 is a Gaussian:

$$f^{(v)}(\boldsymbol{v}) \propto \exp\left(-\lambda v^2\right),$$

where λ is a positive constant. Show that

$$\left\langle v^2 \right\rangle = \frac{3}{2\lambda},$$

and that therefore the average kinetic energy is given by

$$\left\langle \frac{1}{2} m v^2 \right\rangle = \frac{3m}{4\lambda}.$$

1.3.3 Equation of State

We now prove that Maxwell's postulates allow us to derive the equation of state for ideal gases and provide a microscopic interpretation of absolute temperature in terms of kinetic energy.

Let us consider a particle of velocity $\boldsymbol{v} = (v_x, v_y, v_z)$ that, coming from the left, hits a wall parallel to the plane (yz) (see figure 1.1). After the impact, it is driven by velocity $\boldsymbol{v}' = (-v_x, v_y, v_z)$. The change $\Delta \boldsymbol{p}$ in its momentum $\boldsymbol{p}$ is given by $\Delta \boldsymbol{p} = \boldsymbol{p}' - \boldsymbol{p} = m(\boldsymbol{v}' - \boldsymbol{v}) = m(-2v_x, 0, 0)$. The number of impacts of this type that occur in a time interval Δt on a certain region of the wall of area S is equal to the number of particles driven by velocity $\boldsymbol{v}$ that are contained in a box of base equal to S and of height equal to $v_x\, \Delta t$. The volume of this box is equal to $S\, v_x\, \Delta t$, and the number of these particles is equal to $\rho f^{(v)}(\boldsymbol{v})\, v_x\, S\, \Delta t$.

The total momentum $\Delta \boldsymbol{P}$ transmitted from the wall to the gas, during the time interval Δt, is therefore

$$\Delta \boldsymbol{P} = \int_0^{+\infty} \mathrm{d}v_x \int_{-\infty}^{+\infty} \mathrm{d}v_y \int_{-\infty}^{+\infty} \mathrm{d}v_z f^{(v)}(\boldsymbol{v})\, \rho\, S\, \Delta t\, (-2m)\, v_x^2\, \boldsymbol{i}, \tag{1.1}$$

where $\boldsymbol{i} = (1, 0, 0)$ is the versor of the x axis. In this expression, the integral over v_x runs only on the region $v_x > 0$ because only those particles that are moving toward the right contribute to pressure on the wall.

The total force that the wall exercises on the gas is given by $\boldsymbol{F} = -\Delta \boldsymbol{P}/\Delta t$, and therefore the pressure p is given by

$$p = \frac{|\boldsymbol{F}|}{S} = 2m\rho \cdot \frac{1}{2} \langle v_x^2 \rangle = \frac{\rho m}{2\lambda}. \tag{1.2}$$

In this equation, the factor $1/2$ comes from the integration over v_x, which runs only on the region $v_x > 0$. It is well known that the equation of state for perfect gases takes the form

$$pV = nRT, \tag{1.3}$$

where $n = N/N_A$ is the number of moles, T is the absolute temperature, and $R \simeq 8.314\ \mathrm{J\ K^{-1}\ mol^{-1}}$ is the gas constant. By introducing the **Boltzmann constant** k_B,

$$k_B = \frac{R}{N_A} \simeq 1.381 \cdot 10^{-23}\ \mathrm{J\ K^{-1}}, \tag{1.4}$$

and the particle density

$$\rho = \frac{N}{V}, \tag{1.5}$$

eq. (1.3) can be written

$$p = \rho\, k_B T. \tag{1.6}$$

If we compare this expression with eq. (1.2), we obtain the constant λ:

$$\lambda = \frac{m}{2k_B T}. \tag{1.7}$$

The distribution of the component v_x of the velocity, taking into account the normalization condition, is then given by

$$f^{(v_x)}(v) = \left(\frac{m}{2\pi\ k_B T}\right)^{1/2} \exp\left(-\frac{mv^2}{2k_B T}\right). \tag{1.8}$$

Analogous laws hold for the distributions of v_y and v_z. Then the distribution of the speed $v = |\boldsymbol{v}|$ is given by the following expression, known as the **Maxwell distribution**:

$$\phi(v) = \int \mathrm{d}\boldsymbol{v}\, \delta\left(v - |\boldsymbol{v}|\right) f^{(v)}(\boldsymbol{v}) = \left(\frac{m}{2\pi\ k_B T}\right)^{3/2} 4\pi v^2\ \exp\left(-\frac{mv^2}{2\,k_B T}\right), \tag{1.9}$$

where $\delta(x)$ is the **Dirac delta** "function" (cf. sec. A.8), defined by

$$\int \mathrm{d}x f(x)\, \delta(x - x_0) = f(x_0), \tag{1.10}$$

where $f(x)$ is an arbitrary smooth function.

Exercise 1.3 (On Maxwell's distribution)

1. Evaluate the root mean square speed $v_{\mathrm{RMS}} = \sqrt{\langle v^2 \rangle}$ of a gas obeying Maxwell's distribution $\phi(\boldsymbol{v})$ in d dimensions. Give the corresponding numerical value for air in three dimensions at room temperature ($T = 300$ K). Air is made by 2/3 of nitrogen (molecular weight 28) and 1/3 of oxygen (molecular weight 32).

2. Evaluate the value v of the speed in a Maxwell distribution at temperature T for a gas of molecules of mass m in d dimensions that corresponds to the maximal probability density. Compare it with the root mean square speed.

3. Evaluate the probability distribution function for the kinetic energy of a gas following Maxwell's distribution, i.e., the probability density that the kinetic energy of a randomly chosen particle has the value κ:

$$f(\kappa) = \int \mathrm{d}\boldsymbol{v}\, \delta\left(\kappa - \frac{1}{2} m v^2\right) \phi(\boldsymbol{v}).$$

Exercise 1.4 (Gas in a gravity field) Let us consider a gas of particles of mass m, at equilibrium with a uniform temperature T in a gravity field, described by the acceleration $\boldsymbol{g}$ pointing vertically downward. The gas particles obey the Maxwell distribution. Let S_1 and S_2 be two horizontal surfaces, at heights z_1 and $z_2 > z_1$, respectively. We denote by ρ_i, $i \in \{1, 2\}$ the numerical density close to S_i.

By considering the condition that, in a time interval of duration Δt, as many particles originating from S_1 cross S_2 as vice versa, derive a relation between ρ_i and z_i ($i \in \{1, 2\}$). Neglect the effects of collisions between S_1 and S_2.

Exercise 1.5 (Knudsen gas) Let us consider a gas distributed in two containers placed side by side, both at a very small density, at pressures and temperatures (p_1, T_1) and (p_2, T_2). The two containers are connected by a very small opening, such that the thermal equilibrium of the two systems is not perturbed. There is molecular equilibrium when the flux of particles through the opening from container 1 to container 2 is equal to the flux of particles from container 2 to container 1.

1. Show that in this situation there is a simple relation between (p_1, T_1) and (p_2, T_2).

2. Evaluate the small change in (p_i, T_i), $i = 1, 2$, that obtains in a short time interval Δt.

Exercise 1.6 (Drag in a gas) Let us consider an ideal gas made of particles of mass m, with numerical density $\rho = N/V$, at temperature T. Using Maxwell's distribution, evaluate the drag applied by the gas on a small disk of radius R moving at velocity $\boldsymbol{v}$, where $\boldsymbol{v}$ is parallel to the axis of the disk. We assume that R is much smaller than the interatomic distance $\rho^{-1/3}$ and that v is much smaller than the characteristic speed of the particles.

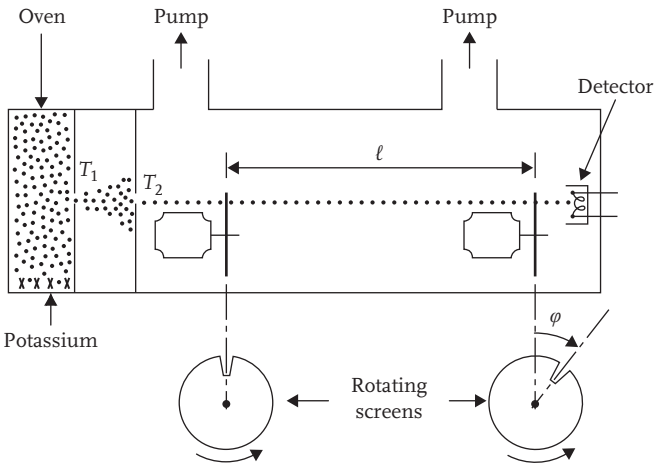


FIGURE 1.2. The Marcus and McFee experiment. Based on [Marc59].

1.3.4 *The Marcus and MacFee experiment*

The Maxwell distribution can be measured directly by means of experiments on the molecular beams. We will follow the work by Marcus and McFee [Marc59]. A diagram of the experiment is given in figure 1.2. Potassium atoms are heated to a fairly high temperature (a few hundred degrees Celsius) in an oven. The oven is equipped with a small opening that allows the beam to escape. A screen with a small hole is placed a bit further away, in order to have a well-aligned beam. In the region traversed by the beam, a vacuum is maintained by a pump. Two rotating screens, set at a distance ℓ from each other, act as velocity selectors. Each is endowed with a narrow gap, and they rotate together with angular velocity ω. The two gaps are out of phase by an angle φ. Therefore, only particles driven by a velocity $v = \ell\omega/\varphi$ will be able to pass through both gaps, hit the detector, and be counted. If we denote the total beam intensity by j_0, and the solid angle by which the detector is seen from the opening by $\mathrm{d}\Omega$, the number of particles driven by a velocity between v and $v + \mathrm{d}v$ that hit the detector in a given unit of time is given by

$$j\,\mathrm{d}v\,\mathrm{d}\Omega = \frac{1}{\mathcal{Z}} j_0 v\,\phi(v)\,\mathrm{d}v\,\mathrm{d}\Omega = \frac{j_0\,m^2}{2\pi\,(2k_\mathrm{B}T)^2}\,v^3\,\exp\left(-\frac{mv^2}{2k_\mathrm{B}T}\right)\,\mathrm{d}v\,\mathrm{d}\Omega, \tag{1.11}$$

where the normalization constant $\mathcal{Z}$ is chosen so that the total particle flux equals j_0.

By varying φ or ω, one can measure the particle flow at various velocities v. We can introduce the variable $\eta = v\sqrt{m/k_\mathrm{B}T}$, thus obtaining a law that is independent of both m and T:

$$\frac{j\,\mathrm{d}v}{j_0\,\mathrm{d}\Omega} = \frac{\eta^3}{2}\exp\left(-\frac{\eta^2}{2}\right)\,\mathrm{d}\eta. \tag{1.12}$$

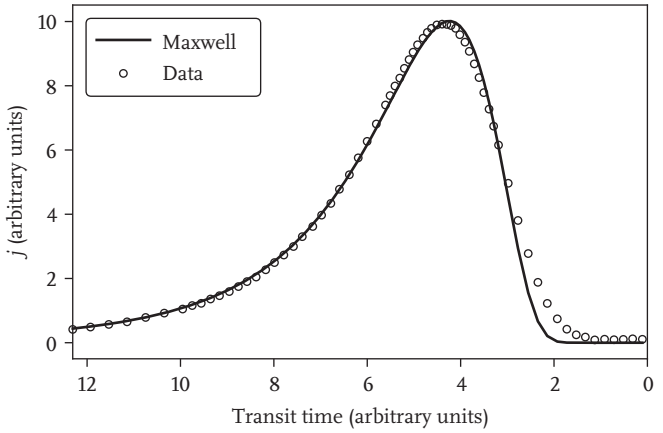


FIGURE 1.3. Particle flux as a function of the transit time between the two disks in the Marcus and McFee experiment at 157°C. The line is a straightforward fit to the Maxwell distribution. The discrepancy at high velocity (small transit time) is due to the fact that the disks do not instantaneously interrupt the particle flow. Based on [Marc59].

This law is well suited to experimental proof, which is shown in figure 1.3. In order to arrive at a quantitative agreement with the results of the experiment, it is necessary to take into account the fact that the disks do not instantaneously interrupt the particle flow and that therefore there are uncertainties in the velocity selection.

1.4 Conclusions

We have therefore been able to show that, if one formulates some statistical hypotheses about the distribution of the velocities of the particles, the equation of state of ideal gases is compatible with the ideal gas model. In the argument we have laid out, there is, however, no proof that the statistical hypotheses concerning the distribution of position and velocity of particles are compatible with the laws of mechanics. In order to prove this compatibility, we need to establish some hypotheses about initial conditions. We will therefore be satisfied if we manage to prove that for "almost all" initial conditions that satisfy certain criteria, the statistical properties of the relevant distributions are "almost always" valid. These considerations will be incomplete, however, since it is not clear how hypotheses that are valid for almost all initial conditions can be relevant to explain the result of a specific experiment, carried out with specific initial conditions.

The statistical postulates are ultimately founded in thermodynamics. S.-K. Ma stressed in [Ma85] that they can be summarized by the relation (discovered by Boltzmann) between thermodynamic entropy and a mechanical quantity: the volume of the accessible phase space. In order to clarify the nature of the entropy and what we mean by volume of accessible phase space, it is necessary to briefly discuss the principles of thermodynamics, which is done in the next chapter.

Recommended Reading

The kinetic theory of gases is treated in greater detail in a large number of works. A brief but careful introduction can be found in K. Huang, *Statistical Mechanics*, New York: Wiley, 1987. A classic reference work is R. C. Tolman, *The Principles of Statistical Mechanics*, Oxford, UK: Oxford University Press, 1938 (reprinted New York: Dover, 1979). The (often surprising) history of the kinetic theory has been retraced by S. G. Brush, *The Kind of Motion We Call Heat*, Amsterdam: North-Holland, 1976.

2 Thermodynamics

> The method of "postulating" what we want has many advantages; they are the same as the advantages of theft over honest toil.
>
> —Lord Russell

2.1 Thermodynamic Systems

In this chapter, I will present a brief review of thermodynamics, following an order that is the reverse of the most common way by which it is introduced. The usual course is inductive: one starts with certain regularities that can be observed in thermal phenomena in order to present them as general principles (the **principles of thermodynamics**); one then proceeds to define entropy and its properties; and one finally applies these properties to the study of the equilibrium of thermodynamic systems. We will instead use a deductive approach: we will postulate the properties of thermodynamic equilibrium—specifically, the existence of the entropy function—in order to then deduce the properties of thermodynamic equilibrium. This approach, which was established by Gibbs in the 1870s ([Gibb73], [Gibb75]), and then developed by various authors (especially L. Tisza in the 1960s [Tisz66]), is the one followed by most modern textbooks, especially Callen's [Call85], which I shall follow fairly closely.

A **thermodynamic system** is a *macroscopic system* whose behavior is identified thanks to *a small and finite* number of quantities—the thermodynamic properties.

The most common example is a given amount of a **simple fluid**—in other words of a physical phase that is *homogeneous and isotropic*—formed by particles that belong to only one chemical species. In this case, the thermodynamic properties are the volume V, the number of particles N, and the internal energy E. In the case of nonsimple fluids (composed of mixtures of particles belonging to several chemical species), one will have to consider the numbers of particles N_1, N_2, and so on of the relevant chemical species.

One can also examine more general thermodynamic properties—for example, in magnetic systems, one considers the total magnetic dipole $\boldsymbol{M}$ of the sample, usually called **magnetization**; in crystals, there are several thermodynamic quantities that characterize its structure and therefore allow one to distinguish between the various crystalline phases.

Thermodynamics is a "miracle": the system has $\sim 10^{23}$ degrees of freedom, and yet it can be well described by a much smaller number of variables. It is clear that we will not be able to foresee all the properties of our system, but for our purposes those few parameters will be sufficient. There is a certain degree of circularity in the definition of thermodynamic parameters, which is resolved by experiment. One considers only a restricted set of manipulations on thermodynamic systems. In practice, one allows them to be put in contact with one another, or one acts upon them by changing a few macroscopic properties such as their volume or the electric or magnetic field in which they are immersed. One then identifies a number of properties such that, if they are known before the manipulation, their values after the manipulation can be predicted. The smallest set of properties that can successfully be predicted in this way can be selected as the basis for a thermodynamic description of the system. Unfortunately, we are not able, at least for the time being, to define a priori, for any given system, the relevant thermodynamic properties that characterize its behavior. In order to do this, we must resort to experiments.

If the state of a thermodynamic system can be fully characterized by the values of the thermodynamic variables, and if these values are invariant over time, one says that the system is in a state of **thermodynamic equilibrium**. Thermodynamic equilibrium, as R. Feynman [Feyn72] says, occurs when all fast processes have already occurred, while the slow ones have yet to take place. Clearly, the distinction between fast and slow processes depends on the observation time $\mathcal{T}$ that is being considered. This property has been emphasized by S.-K. Ma [Ma85, p. 3]:

> If we pour some boiling water from a thermos flask into a tea cup, after a few seconds the water in the cup becomes an unchanging state at rest, and this is an equilibrium state. The volume of the water in the cup can be measured from the height of the water and the cross section area of the cup. The temperature of the water can be measured by a thermometer. Within one or two seconds, these measured quantities will not change significantly, and thus during this observation time the water is in equilibrium. If the time is too long, the temperature will obviously change and the water cannot be regarded as being in an equilibrium state. After an hour, the temperature of the water will be equal to the room temperature. If the room temperature is constant for several hours, then the temperature of the water will accordingly remain constant. Therefore, if the observation time is within this range of several hours, the cup of water can be regarded as being in equilibrium. However, water molecules will continually evaporate from the water surface. In several hours the volume of the water will not change significantly, but if the time is long enough, such as two or three days, all the water will evaporate. So if the observation period is over two or three days, this cup of water cannot be regarded as being in an equilibrium state. After a few days, when all the water has evaporated, then again this empty cup can be regarded as being in an equilibrium state. However, strictly speaking, even this cup is not in an absolute unchanging state, because the molecules of the cup can evaporate, although in a few years' time there will be no change observed.

These arguments show that a system can be shown to be in equilibrium if the observation time is fairly short, while it is no longer possible to consider it in equilibrium for longer observation times. A more curious situation, but one that often occurs, is that the

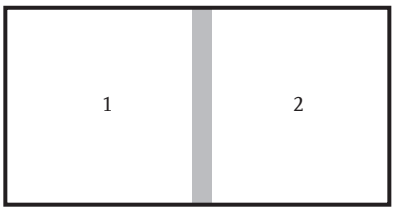


FIGURE 2.1. Systems in contact. The gray area represents the region in which the interaction between the two systems takes place.

same system can be considered in equilibrium, but with different properties for different observation times. For example, a block of tar in the course of a period of observation lasting several minutes behaves like an elastic solid. If, however, it is left alone in a container, after several hours have elapsed, it will take the shape of the container like any liquid. It is therefore possible to consider it as if it were a solid for brief periods of time and as a liquid over longer periods of time. Similarly, basalt is commonly considered a solid, but in the context of continental drift (in other words, for observation times lasting several million years), it is possible to consider it a liquid on which the continents float. We shall see that analogous phenomena also occur for much less exotic systems, like gaseous hydrogen. There are some very "stubborn" systems where one cannot forget this fact because the set of thermodynamic properties changes too rapidly with variations in observation times—this is the case for all types of glass, which in a certain sense represents the frontier of thermodynamic description.

2.2 Extensive Variables

Let us consider two thermodynamic systems, 1 and 2, that can be made to interact with one another. Variables like the volume V, the number of particles N, and the internal energy E, whose values (relative to the total system) are equal to the sum of the values they assume in the single systems, are called **additive** or **extensive**.

Let us note that, strictly speaking, internal energy is *not* extensive. In reality if we put two systems in contact with each other, the total energy of the system $1 \cup 2$ will also receive the contribution of the interaction between the two systems 1 and 2. More generally, we will therefore have

$$E^{(1\cup 2)} = E^{(1)} + E^{(2)} + E_{\text{int}}, \tag{2.1}$$

where the indexes (1) and (2) refer to the two subsystems, and E_{int} is the interaction energy. If the particles making up the two systems have only short-range interactions, however, the effect of the interaction between the two systems will only be felt on the surface in which they are in contact, and for a limited thickness on either side (see fig. 2.1). Therefore, while one might expect $E^{(1)}$ and $E^{(2)}$ to be proportional to the volume of their respective systems, E_{int} will be proportional to their surface. For large systems, as long as they occupy well-defined portions of space, surface grows in size much more slowly than volume, so that ultimately the contribution of E_{int} is shown to be negligible (see figure 2.1). So if a container holds water and oil, the interaction energy between the water and the oil is proportional

to the container's cross section, while the energy of the water (or oil) is proportional to its respective volume. In some situations, however, it is possible to obtain structures in which, for instance, oil forms droplets $\sim 50\,\text{nm}$ in diameter, which are dispersed in the water and are in equilibrium with it. These structures are called **microemulsions**. In this case, the interaction energy between water and oil becomes proportional to the container's volume, because within each fairly large element of the system's volume a certain amount of interface is present. In this case, it is not possible to consider the system as made of water and oil set side by side, but it is more practical to consider it a thermodynamically homogeneous system.

It is not possible to neglect the interaction energy in the presence of long-range interactions, as, for instance, in the case of electrical, magnetic, or gravitational fields. It is sometimes possible to consider the effects of these interactions explicitly; one therefore considers, for example, the gravitational interaction energy between two bodies as *mechanical* (and not *internal*) energy.

The **fundamental hypothesis of thermodynamics** is that it should be possible to characterize the state of a thermodynamic system by specifying the values of a certain set $(X_0, X_1, \ldots, X_r)$ of *extensive* variables. For example, X_0 could be the internal energy E, X_1 the number of particles N, X_2 the volume V of the system, and so on.

2.3 The Central Problem of Thermodynamics

Classical thermodynamics deals with *equilibrium states* of *thermodynamic systems*. The central problem of thermodynamics can be formulated as follows:

> *Given the initial state of equilibrium of several thermodynamic systems that are allowed to interact, determine the final thermodynamic state of equilibrium.*

The interaction between thermodynamic systems is usually represented by idealized walls that allow the passage of one (or more) extensive quantities from one system to the other. Among the various possibilities, the following are usually considered:

Thermally conductive (diathermal) walls: These allow the exchange of energy, but not of volume or particles.

Semipermeable walls: These allow the exchange of particles belonging to a given chemical species (and consequently also of energy).

Adiabatic walls: These prevent the exchange of energy.

Figure 2.2 shows a generic diagram of a thermodynamic experiment. Systems 1 and 2 are each initially in a state of thermodynamic equilibrium characterized by their respective values of the extensive thermodynamic variables $(X_0, X_1, \ldots, X_r)$, or, in other words, by $(X_{0\text{in}}^{(1)}, X_{1\text{in}}^{(1)}, \ldots, X_{r\text{in}}^{(1)})$ and $(X_{0\text{in}}^{(2)}, X_{1\text{in}}^{(2)}, \ldots, X_{r\text{in}}^{(2)})$. They are then put in contact—that is, one allows the exchange of some of these extensive quantities between one system and the other to occur.

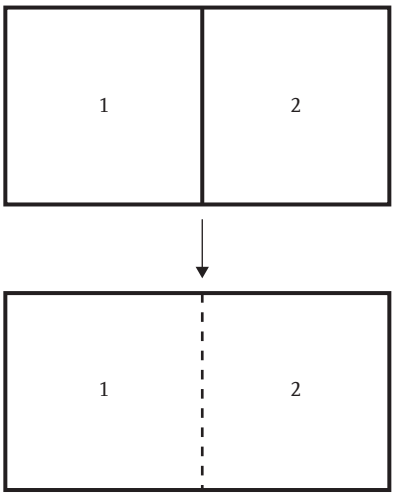


FIGURE 2.2. Generic diagram of a thermodynamic experiment. The wall separating systems 1 and 2 is either removed or relaxed, in order to allow the exchange of some extensive quantities.

After a certain amount of time has elapsed (whose duration is not of concern to thermodynamics), a new state of equilibrium is reached, characterized by the values $(X^{(1)}_{0\text{fin}}, X^{(1)}_{1\text{fin}}, \ldots, X^{(1)}_{r\text{fin}})$ and $(X^{(2)}_{0\text{fin}}, X^{(2)}_{1\text{fin}}, \ldots, X^{(2)}_{r\text{fin}})$ for system 1 and system 2, respectively. The problem is to characterize this state of equilibrium.

The *possible* states of equilibrium (compatible with the constraints and the initial conditions) form the space of **virtual states**. The initial state is obviously a (specific) virtual state. The central problem of thermodynamics can be thus restated as follows:

Characterize the actual state of equilibrium among all virtual states.

2.4 Entropy

The solution of the central problem of thermodynamics is provided by the following **entropy postulate**:

*There exists a function S of the extensive variables $(X_0, X_1, \ldots, X_r)$, called **entropy**, that assumes the maximum value for a state of equilibrium among all virtual states and that possesses the following properties:*

Extensivity: If 1 and 2 are thermodynamic systems, then

$$S^{(1\cup 2)} = S^{(1)} + S^{(2)}. \tag{2.2}$$

Concavity: If $X^{(1)} = (X^{(1)}_0, X^{(1)}_1, \ldots, X^{(1)}_r)$ and $X^{(2)} = (X^{(2)}_0, X^{(2)}_1, \ldots, X^{(1)}_r)$ are two thermodynamic states of the same system, then for any α between 0 and 1, one has

$$S\left((1-\alpha)X^{(1)} + \alpha X^{(2)}\right) \geq (1-\alpha)S(X^{(1)}) + \alpha S\left(X^{(2)}\right). \tag{2.3}$$

In other words, $S(X)$ is **concave**, or, equivalently, $-S(X)$ is a **convex** function of the extensive variables. The properties of convex functions are discussed in sec. B.1.

There, one shows that eq. (2.3) implies

$$\sum_{i=0}^{r} \left.\frac{\partial S}{\partial X_i}\right|_{X^{(1)}} \left(X_i^{(2)} - X_i^{(1)}\right) \geq S(X^{(2)}) - S\left(X^{(1)}\right), \tag{2.4}$$

which expresses the fact that the surface $S(X_0, X_1, \ldots, X_r)$ lies always *below* the tangent plane in one of its points.

Monotonicity: $S(E, X_1, \ldots, X_r)$ is a monotonically increasing function of the internal energy E:

$$\left.\frac{\partial S}{\partial E}\right)_{X_1,\ldots,X_r} = \frac{1}{T} \geq 0. \tag{2.5}$$

There can be some cases, as we will see, in which this condition is violated—these cases are fairly special, because they correspond to the description of certain relatively isolated subsystems, in which the energy spectrum has a finite upper bound. In ordinary cases, however, in which energy does not admit a finite upper bound, this condition is always satisfied.

The entropy postulate allows one to solve the central problem of thermodynamics, by transforming it into a **constrained extremum** problem:

The equilibrium state corresponds to the maximum entropy compatible with the constraints.

Since the initial state X^{in} of an isolated system is one particular virtual state, it follows that the value of the entropy of the equilibrium state X^{eq} reached after some constraints are relaxed cannot be smaller than its initial value:

$$S(X^{\text{eq}}) \geq S(X^{\text{in}}). \tag{2.6}$$

This is a formulation of the **second principle of thermodynamics**.

In the next section, we examine some simple applications of this postulate.

2.5 Simple Problems

2.5.1 *Thermal Contact*

Let us consider two systems, 1 and 2, that are in contact by means of a thermally conductive wall—they can therefore exchange internal energy, but not other extensive quantities. The virtual state space is therefore defined by the relations:

$$E^{(1)} + E^{(2)} = X_0^{(1)} + X_0^{(2)} = E = \text{const.}, \tag{2.7}$$

$$X_i^{(1)} = \text{const.}, \qquad i = 1, 2 \ldots, r, \tag{2.8}$$

$$X_i^{(2)} = \text{const.}, \qquad i = 1, 2, \ldots, r. \tag{2.9}$$

Let us look for the maximum of S as a function, for example, of $E^{(1)}$. By eq. (2.7) the partial derivative of the total entropy with respect to $E^{(1)}$ is given by

$$\begin{aligned}\frac{\partial S}{\partial E^{(1)}} &= \frac{\partial}{\partial E^{(1)}}\left[S^{(1)}(E^{(1)},\dots,X_r^{(1)}) + S^{(2)}(E-E^{(1)},\dots,X_r^{(2)})\right] \\ &= \left.\frac{\partial S^{(1)}}{\partial E^{(1)}}\right|_{E^{(1)}} - \left.\frac{\partial S^{(2)}}{\partial E^{(2)}}\right|_{E-E^{(1)}}.\end{aligned} \tag{2.10}$$

(Some useful properties of partial derivatives are listed in sec. C.) The derivatives are understood to be taken with $X_i^{(1),(2)}$ fixed for $i=1,\dots,r$. We denote by $E_{\text{eq}}^{(1)}$ the value of $E^{(1)}$ at equilibrium. Since $\partial S/\partial E^{(1)}$ vanishes at equilibrium, we have

$$\left.\frac{\partial S^{(1)}}{\partial E^{(1)}}\right|_{E_{\text{eq}}^{(1)}} = \left.\frac{\partial S^{(2)}}{\partial E^{(2)}}\right|_{E_{\text{eq}}^{(2)}}, \tag{2.11}$$

where $E_{\text{eq}}^{(2)} = E - E_{\text{eq}}^{(1)}$ is the equilibrium value of $E^{(2)}$.

It is useful to ask ourselves which of the two subsystems has released internal energy. Denoting the initial values of E with E_{in}, we have, since the initial state is a virtual state,

$$S^{(1)}(E_{\text{eq}}^{(1)}) + S^{(2)}(E_{\text{eq}}^{(2)}) \geq S^{(1)}(E_{\text{in}}^{(1)}) + S^{(2)}(E_{\text{in}}^{(2)}), \tag{2.12}$$

and therefore

$$S^{(1)}(E_{\text{eq}}^{(1)}) - S^{(1)}(E_{\text{in}}^{(1)}) + S^{(2)}(E-E_{\text{eq}}^{(1)}) - S^{(2)}(E_{\text{in}}^{(2)}) \geq 0. \tag{2.13}$$

Due to the concavity of entropy, one has

$$S(E_{\text{eq}}) - S(E_{\text{in}}) \leq \left.\frac{\partial S}{\partial E}\right|_{E_{\text{in}}} (E_{\text{eq}} - E_{\text{in}}). \tag{2.14}$$

We thus obtain

$$\left[\left.\frac{\partial S^{(1)}}{\partial E^{(1)}}\right|_{E_{\text{in}}^{(1)}} - \left.\frac{\partial S^{(2)}}{\partial E^{(2)}}\right|_{E_{\text{in}}^{(2)}}\right]\left(E_{\text{eq}}^{(1)} - E_{\text{in}}^{(1)}\right) \geq 0. \tag{2.15}$$

Energy therefore shifts toward the system with the higher values of $\partial S/\partial E$.

Let us introduce the quantity

$$T = \left(\frac{\partial S}{\partial E}\right)^{-1}, \tag{2.16}$$

where the derivative is taken with all other extensive quantities kept constant. According to our hypotheses, this quantity is *positive*. We thus obtain the following results:

- At equilibrium, T is the same in all subsystems that are in reciprocal contact by means of thermally conductive walls.
- In order to reach equilibrium, energy shifts from systems with higher values of T toward systems with lower values of T.

These properties tie T to empirical temperatures. And in fact, it is possible to deduce from them that T is a monotonically increasing function of temperature. We will see later that it is possible to identify T with the *absolute temperature*.

2.5.2 A Thermally Conductive and Mobile Wall

In this case, the two systems can also exchange volume V, in addition to internal energy E. The two equilibrium conditions are therefore:

1. The condition of maximum entropy relative to exchange of internal energy, as already seen:

$$\left.\frac{\partial S^{(1)}}{\partial E^{(1)}}\right|_{E^{(1)}_{\text{eq}}} = \left.\frac{\partial S^{(2)}}{\partial E^{(2)}}\right|_{E^{(2)}_{\text{eq}}}. \tag{2.17}$$

2. The condition of maximum entropy relative to exchange of volume, which can be obtained in an analogous fashion:

$$\left.\frac{\partial S^{(1)}}{\partial V^{(1)}}\right|_{V^{(1)}_{\text{eq}}} = \left.\frac{\partial S^{(2)}}{\partial V^{(2)}}\right|_{V^{(2)}_{\text{eq}}}. \tag{2.18}$$

By introducing the quantity p, defined by the equation

$$\frac{p}{T} = \frac{\partial S}{\partial V}, \tag{2.19}$$

where the partial derivative is taken by keeping the other extensive quantities constant, these conditions can be expressed as follows:

$$T^{(1)} = T^{(2)}, \tag{2.20}$$

$$p^{(1)} = p^{(2)}. \tag{2.21}$$

We show later that the quantity p can be identified with **pressure**.

One can easily prove that between two systems, both initially at the same temperature, volume is initially released by the system in which pressure is lower into the system in which pressure is higher. In fact, because of the principle of maximum entropy, one has

$$S^{(1)}(E^{(1)}_{\text{eq}}, V^{(1)}_{\text{eq}}) - S^{(1)}(E^{(1)}_{\text{in}}, V^{(1)}_{\text{in}}) + S^{(2)}(E^{(2)}_{\text{eq}}, V^{(2)}_{\text{eq}}) - S^{(2)}(E^{(2)}_{\text{in}}, V^{(2)}_{\text{in}}) \geq 0, \tag{2.22}$$

while because of concavity, for each system one has

$$S(E_{\text{eq}}, V_{\text{eq}}) - S(E_{\text{in}}, V_{\text{in}}) \leq \left.\frac{\partial S}{\partial E}\right|_{E_{\text{in}}} (E_{\text{eq}} - E_{\text{in}}) + \left.\frac{\partial S}{\partial V}\right|_{V_{\text{in}}} (V_{\text{eq}} - V_{\text{in}}). \tag{2.23}$$

When we apply this relation to our case, we obtain

$$\begin{aligned}&\left[\left.\frac{\partial S^{(1)}}{\partial E^{(1)}}\right|_{E^{(1)}_{\text{in}}} - \left.\frac{\partial S^{(2)}}{\partial E^{(2)}}\right|_{E^{(2)}_{\text{in}}}\right]\left(E^{(1)}_{\text{eq}} - E^{(1)}_{\text{in}}\right)\\ &\quad + \left[\left.\frac{\partial S^{(1)}}{\partial V^{(1)}}\right|_{V^{(1)}_{\text{in}}} - \left.\frac{\partial S^{(2)}}{\partial V^{(2)}}\right|_{V^{(2)}_{\text{in}}}\right]\left(V^{(1)}_{\text{eq}} - V^{(1)}_{\text{in}}\right) \geq 0.\end{aligned} \tag{2.24}$$

The first term vanishes because the initial temperatures in the two systems are equal. We thus obtain:

$$\left[\left. \frac{\partial S^{(1)}}{\partial V^{(1)}} \right|_{V_{\text{in}}^{(1)}} - \left. \frac{\partial S^{(2)}}{\partial V^{(2)}} \right|_{V_{\text{in}}^{(2)}} \right] \left(V_{\text{eq}}^{(1)} - V_{\text{in}}^{(1)} \right) \geq 0. \tag{2.25}$$

Since the temperatures T in the two systems are the same, this relation implies that

$$\left(p^{(1)} - p^{(2)} \right) \left(V_{\text{eq}}^{(1)} - V_{\text{in}}^{(1)} \right) \geq 0. \tag{2.26}$$

The apparently similar case of two systems separated by a wall that is both mobile (allowing the exchange of volume) and adiabatic (preventing the free exchange of internal energy), cannot be solved within thermodynamics alone, as we discuss later.

2.5.3 *A Semipermeable Wall*

Let us consider a system composed of several chemical species, and let us introduce the numbers of molecules, $N_1, N_2, \ldots, N_s$, belonging to the chemical species that constitute it as part of the thermodynamic variables. Let us suppose that two systems of this type are separated by a semipermeable wall, that is, a wall that only allows the kth chemical species to pass. Clearly, it is impossible for the exchange of molecules to occur without an exchange of energy. The equilibrium conditions will therefore be

$$T^{(1)} = T^{(2)}, \tag{2.27}$$

$$\frac{\partial S^{(1)}}{\partial N_k^{(1)}} = \frac{\partial S^{(2)}}{\partial N_k^{(2)}}. \tag{2.28}$$

We define the **chemical potential** μ_k of the kth chemical species by

$$\frac{\mu_k}{T} = -\frac{\partial S}{\partial N_k}. \tag{2.29}$$

The conditions of equilibrium can be then recast in the form

$$T^{(1)} = T^{(2)}, \tag{2.30}$$

$$\mu_k^{(1)} = \mu_k^{(2)}. \tag{2.31}$$

Exercise 2.1 Consider two systems, initially at the same temperature, that are in contact by means of a semipermeable wall that allows for the passage of particles of the chemical species k. Show that, in this case, the particles flow from the system with a larger value of μ_k to that with a smaller value.

2.6 Heat and Work

The consistency of thermodynamics requires that the internal energy be a function of the thermodynamic variables. Some of these variables are mechanical in nature, in other

words, they are not thermodynamic—for example, the volume V, the number of particles N, the total magnetic moment $\boldsymbol{M}$, and so on.

From mechanics (and from electromagnetism), we can derive an expression for the infinitesimal mechanical **work** performed on the system by varying the extensive quantities. One usually adopts a sign convention according to which the work W is considered positive if the system performs work on the environment, and negative if the opposite is true. Following this convention, the expression of the infinitesimal work $đW$ is given by

$$đW = -\sum_{i=1}^{r} f_i \, \mathrm{d}X_i. \tag{2.32}$$

The symbol đ reminds us that $đW$ is not an "exact" differential—it is not the differential of a function, because the integral $\int đW$ depends on the particular path being considered, and not only on the initial and final states.

The quantities f_i that appear in this expression are called **generalized forces**. If the system is thermally isolated, it can exchange energy with the outside only by means of mechanical work. In this case, $đW$ is equal to the opposite of the infinitesimal variation of the internal energy of the system:

$$đW = -\mathrm{d}E = -\sum_{i=1}^{r} f_i \, \mathrm{d}X_i. \tag{2.33}$$

It is well known that although *perfectly adiabatic* walls that completely prevent the exchange of energy in the form of heat do not exist, using a series of devices, it is possible to approximate this ideal condition.

If the walls are not adiabatic, we will have in general

$$\mathrm{d}E = đQ - đW, \tag{2.34}$$

which represents the definition of the infinitesimal **heat** $đQ$ exchanged in the process. As highlighted by the notation, $đQ$ is not an exact differential, either. The fact that $\mathrm{d}E$, expressed in this way, is an exact differential is one formulation of the **first principle of thermodynamics**.

2.6.1 Temperature

Let us now show that the quantity T, defined by eq. (2.16), is indeed the *absolute temperature* that one encounters in elementary thermodynamics. We rewrite (2.16) as follows:

$$\left.\frac{\partial S}{\partial E}\right)_X = \frac{1}{T}, \tag{2.35}$$

where $X = (X_1, \ldots, X_r)$ denotes the extensive thermodynamic variables other than the internal energy.

Let us review the intuitive properties of the temperature:

Thermal equilibrium: Two bodies in thermal equilibrium have the same temperature.

Heat flux: Heat flows spontaneously from bodies at higher temperatures to bodies at lower temperatures.

As we have seen in the discussion of eqs. (2.11) and (2.15), these properties are also satisfied by our definition of T. Thus T is a monotonically increasing function of the thermodynamical temperature.

In order to distinguish momentarily between two definitions of temperature, let us denote by Θ the absolute temperature defined in elementary thermodynamics. This quantity is identified by Carnot's theorem on the efficiency of thermal engines:

The maximal efficiency of a thermal engine working between the absolute temperatures Θ_1 and Θ_2 (with $\Theta_1 > \Theta_2$) is given by $\eta = 1 - \Theta_1/\Theta_2$.

We will now show that $\Theta \propto T$, where T is "our" temperature. It is then clear that it is possible to choose the entropy scale in such a way as to put the proportionality constant equal to 1.

A thermal engine works by extracting energy from a reservoir at a given temperature T_1, converting a part of this energy into mechanical work on the environment by means of a device, and releasing the remainder to a reservoir at a temperature T_2. The device acts cyclically, that is, it is found in its initial state at the end of the procedure. A heat reservoir is a system that is assumed to be so large that it can exchange an arbitrary amount of energy without changing its temperature: that is, for which the temperature $T = [\partial S/\partial E)_X]^{-1}$ is independent of E. The whole compound system is enclosed in a container that allows it to exchange energy with the environment only in a purely mechanical way, for example, by raising or lowering a macroscopic weight.

Let the system evolve from an initial equilibrium condition, in which the first heat reservoir has internal energy E_1, the second has internal energy E_2, and the thermal engine is in some equilibrium state, to a final equilibrium state in which the first heat reservoir has internal energy E_1' (and the same temperature T_1, since it is a heat reservoir), the second has E_2', the weight has its potential energy changed by the amount $W = (E_1 + E_2) - (E_1' + E_2')$, and the thermal engine is back to its initial state. By definition, the efficiency of the engine is given by $\eta = W/(E_1 - E_1')$, that is, by the ratio of the work performed on the environment to the energy extracted as heat from the high-temperature reservoir.

In a transformation of this kind, the total entropy of the compound system cannot become smaller. Thus, we have

$$S^{(1)}(E_1) + S^{(2)}(E_2) \leq S^{(1)}(E_1') + S^{(2)}(E_2'), \tag{2.36}$$

where $S^{(i)}(E)$ is the entropy of reservoir i ($i = 1, 2$) expressed as a function of its internal energy $E^{(i)}$. This equation holds because the device is back to its initial state. Now, for our heat reservoirs, we have

$$S^{(i)}(E_i') = S^{(i)}(E_i) + \frac{E_i' - E_i}{T_i}, \qquad i = 1, 2. \tag{2.37}$$

Thus we have

$$S^{(1)}(E_1) - S^{(1)}(E_1') = \frac{E_1 - E_1'}{T_1} \leq S^{(2)}(E_2') - S^{(1)}(E_2) = \frac{E_2' - E_2}{T_2}. \tag{2.38}$$

Therefore

$$W = (E_1 - E_1') + (E_2 - E_2') \leq (E_1 - E_1')\left(1 - \frac{T_2}{T_1}\right). \tag{2.39}$$

Comparing this expression with the maximum efficiency evaluated in elementary thermodynamics, we conclude that

$$\frac{\Theta_2}{\Theta_1} = \frac{T_2}{T_1}, \tag{2.40}$$

which implies that $T \propto \Theta$. The proportionality can be chosen so that $T = \Theta$, where the scale of Θ is fixed by the appropriate convention, which we discuss later. We can thus conclude that

$$T = \left[\left(\frac{\partial S}{\partial E}\right)_X\right]^{-1} \tag{2.41}$$

is the absolute temperature.

Consider a transformation in which the system exchanges energy with the environment purely in the form of heat, that is, by keeping all extensive quantities but E constant. We then have $\mathrm{d}E = đQ$, where $đQ$ is the heat released to the system. We thus obtain

$$\mathrm{d}S = \left(\frac{\partial S}{\partial E}\right)_X đQ = \frac{đQ}{T}. \tag{2.42}$$

This is the well-known relation between the heat exchanged and the entropy change that appears in elementary thermodynamics.

2.6.2 Pressure

Let us consider an infinitesimal variation of V while keeping E constant. In this case, mechanics tells us that the work $đW$ performed by the system is given by the product of the pressure P with the infinitesimal variation in volume $\mathrm{d}V$:

$$đW = P\mathrm{d}V. \tag{2.43}$$

On the other hand, since $\mathrm{d}E = 0$, one has

$$đQ = T\mathrm{d}S = đW = P\mathrm{d}V. \tag{2.44}$$

One has therefore

$$\left(\frac{\partial S}{\partial V}\right)_{E,X_2,\ldots} = \frac{P}{T}. \tag{2.45}$$

This allows us to identify the pressure P with the quantity p we defined previously.

As an application of this relation, let us consider two systems separated by an adiabatic and mobile wall, such that the resulting complex system cannot exchange either volume or energy with the environment. Then the internal energy of both systems can change only via the mechanical work associated with the change in their respective volumes.

$$\mathrm{d}E^{(1)} = -p^{(1)}\mathrm{d}V^{(1)} = -\mathrm{d}E^{(2)} = p^{(2)}\mathrm{d}V^{(2)}. \tag{2.46}$$

On the other hand, the total volume is conserved:

$$\mathrm{d}V^{(1)} = -\mathrm{d}V^{(2)}. \tag{2.47}$$

These two equations are in general incompatible, unless $p^{(1)} = p^{(2)}$. Moreover, for an infinitesimal shift of an adiabatic wall, one has

$$\mathrm{d}S = \frac{\partial S}{\partial E}\bigg)_V \mathrm{d}E + \frac{\partial S}{\partial V}\bigg)_E \mathrm{d}V = \left(-\frac{p}{T} + \frac{p}{T}\right)\mathrm{d}V = 0 \tag{2.48}$$

for every value of $\mathrm{d}V$. Consequently, given the hypotheses we made, the system cannot reach equilibrium, unless it is already there, and vice versa, equilibrium conditions do not determine $V^{(1)}$ (or $V^{(2)}$).

Suppose we move such an adiabatic wall out of its equilibrium position. Then the pressure disequilibrium will force the wall to oscillate. Clearly, sooner or later the oscillation will cease, since friction dissipates energy, and therefore makes eq. (2.46) invalid. The final condition of equilibrium will, however, be determined by the details of energy dissipation due to friction.

2.6.3 Chemical Potential

The case of chemical potential μ_k is rather special. If we attempt to vary the number of particles of certain chemical species that belong to a specific system, we must necessarily allow a certain unchecked exchange of energy between the system and the surrounding environment to occur, and therefore a certain exchange of heat. Consequently, the work necessary to vary the number of moles of a chemical species within a system is not, strictly speaking, of a mechanical nature, and, as a result, the fundamental relation that defines chemical potential is the one that describes the variation of *entropy* as the number of particles varies:

$$\frac{\partial S}{\partial N_k}\bigg)_{E,X_2,\ldots} = -\frac{\mu_k}{T}. \tag{2.49}$$

In order to interpret this expression, let us imagine that we have available a semipermeable membrane that is impermeable to the chemical species k but allows the flow of the other species. Then we can build up an experiment in which the number of moles of species k present within a certain region is changed by mechanical means. Suppose, for example, that the system is made of a cylinder contained in an adiabatic enclosure. The cylinder contains a solution of two chemical species—we call the greater component the solvent and the smaller component the solute—and is divided in two regions by a piston fitted

with a membrane that is permeable to the solvent, but not to the solute. Such a membrane is called **semipermeable**. We suppose that the solute is present in only one of these regions, of volume V. Initially, we have N molecules of solute within the volume V, with a concentration $\rho = N/V$. We can now move the piston slowly, until it delimits a region of volume $V - \mathrm{d}V$. We assume that the transformation is made so slowly that the system can be considered to be in equilibrium at each instant. In this case, the entropy of the total system (which is contained in an adiabatic enclosure) remains constant. Therefore the transformation is **reversible**, that is, it can be performed in the opposite direction, from the final to the initial state. By this transformation, the concentration of the solvent in the enclosed region does not change, but that of the solute does. Thus we must perform some work on the system in order to change its solute concentration. Therefore there is a pressure p, called the **osmotic pressure**, acting on the piston fitted with the semipermeable membrane.

We see later how we can obtain an explicit relation between the mechanical work exercised by the piston and the difference in chemical potential between the two regions separated by the semipermeable membrane.

2.7 Adiabatic Transformations and the Axiomatic Foundation of the Entropy Principle

It has been recently realized that the considerations that we described earlier provide a way to establish the existence of the entropy function and of its basic properties on an intellectually and intuitively appealing axiomatic basis. We have considered a transformation in which a thermodynamic system is allowed to exchange energy with the environment in a purely mechanical way, and we have exploited the fact that, in this situation, the only allowed transformations are the ones in which the total entropy is not reduced. One can reverse this logic. Following Lieb and Yngvason [Lieb98], let us define **adiabatic transformations** on a thermodynamic system (which may itself be made up of several subsystems, like the heat reservoirs we have considered) as those allowing energy exchange with the environment in a purely mechanical way. It is possible for the system to interact with an arbitrarily complex auxiliary sytem (such as the thermal engine we have introduced) provided that the auxiliary system is found, at the end of the transformation, in its initial state. On the other hand, no hypothesis is posited about the pace of the transformation itself, which can be arbitrarily fast or violent. It is assumed only that the initial and the final states are both thermodynamical equilibrium states.

Let us consider such a system, and let $X = (X_0, \ldots, X_r)$ identify its equilibrium states. It is then possible to introduce a relation $\prec$ among these states. If starting from it is possible to reach X' via an adiabatic transformation, X one writes

$$X \prec X', \tag{2.50}$$

and says that X precedes X', or one writes $X' \succ X$ and says that X' succeeds X. It is quite clear that $\prec$ is a reflexive and transitive relation, but is in general not symmetric. Indeed,

as emphasized, for example, by Carathéodory [Cara09], the second principle of thermodynamics may be traced to the fact that for any equilibrium state X of a thermodynamical system, there are states that cannot be reached from it via an adiabatic transformation, while X can be reached from them.

Let us call two states X and X' **adiabatically equivalent**, and write $X \overset{A}{\sim} X'$ if $X \prec X'$ and $X' \prec X$. Then $\overset{A}{\sim}$ is an equivalence relation. If $X \prec X'$, but $X' \not\prec X$, we will write $X \prec\!\!\prec X'$. Given two states X and X', if either $X \prec X'$ or $X' \prec X$, we say that the two states are **comparable**.

It is then possible to show that the existence, extensivity, and concavity of the entropy function $S(X)$ follow from a few simple and physically appealing properties of the relation $\prec$:

Reflexivity: Since $X \overset{A}{\sim} X$, $\forall X$, one has $X \prec X$, $\forall X$.

Transitivity: If $X \prec X'$ and $X' \prec X''$, then $X \prec X''$.

Consistency: If $X \prec X'$ and $Y \prec Y'$, then $(X, Y) \prec (X', Y')$.

Scale invariance: If $\lambda > 0$ and $X \prec X'$, then $\lambda X \prec \lambda X'$. This identifies the variables appearing in X as extensive variables, so that the system described by λX is λ times larger than the system described by X.

Splitting and recombination: If $0 < \lambda < 1$, then $X \overset{A}{\sim} (\lambda X, (1-\lambda)X)$. Thus the system described by X can be adiabatically split into two smaller but similar systems, described by λX and $(1-\lambda)X$, respectively, and vice versa.

Stability: If $(X, \epsilon Z) \prec (X', \epsilon Z')$ for some Z, Z' and for a sequence of ϵ that tends to zero, then $X \prec X'$. This means that the set of allowed adiabatic processes cannot be changed by coupling the system to an arbitrarily small external system.

Then, it can be shown that there is a function $S(X)$ such that if $X \overset{A}{\sim} X'$, then $S(X) = S(X')$, and if $X \prec\!\!\prec X'$, then $S(X) < S(X')$. Moreover, if X and Y are the states of two systems and (X, Y) denotes the corresponding state of the compound system, one has

$$S(X, Y) = S(X) + S(Y), \tag{2.51}$$

that is, the function $S(X)$ is additive, and for any $\lambda > 0$ and each state X one has

$$S(\lambda X) = \lambda S(X), \tag{2.52}$$

that is, $S(X)$ is extensive.

Let us consider a set Γ of states such that all states in Γ are mutually comparable, that is, such that if both X and X' belong to Γ, then either $X \prec X'$ or $X' \prec X$ (or both). One can then show that all states of the form $((1-\lambda)X^0, \lambda X^1)$, where $0 \leq \lambda \leq 1$ and both X^0 and X^1 belong to Γ, also belong to Γ. It is then possible to give an explicit expression of the entropy function in Γ. One considers two reference states, X^0 and X^1, such that $X^0 \prec\!\!\prec X^1$. (If it is not possible to find two such states, then $S(X)$ is the constant function.) Then, for

any state X comparable with X^0 and X^1, one sets

$$S(X) = \sup\{\lambda : ((1-\lambda)X^0, \lambda X^1) \prec X\}. \tag{2.53}$$

Here we extend the range of λ to values negative or larger than 1, by introducing the conventions that $(X, -Y) \prec X'$ means $X' \prec X$ and that $(X, 0\ Y)$ means X.

This construction recalls a definition by Laplace and Lavoisier in which heat was measured by the amount of ice a body could melt. In our case, the entropy of a state X is measured by the largest amount of X^1 that can be adiabatically transformed into X with the help of X^0.

The entropy function defined in this way is unique up to affine equivalence, that is, $S(X) \longrightarrow aS(X) + b$, with $a > 0$. In order to define a consistent entropy scale for a set of compound systems, it appears that one has to postulate that all the corresponding states are comparable. If this is assumed, a consistent entropy function for all comparable states of compound systems can be defined. Indeed, Lieb and Yngvason [Lieb98] have shown that this comparability hypothesis follows from a few additional axioms concerning the notion of thermal equilibrium and thermal contact, out of which the notion of temperature can be abstracted.

Following this approach, it is possible to establish the basic notions of thermodynamics, and in particular the entropy principle, on a sounder and intuitively appealing basis. A detailed explanation can be found in [Lieb98].

2.8 The Fundamental Equation

It follows from the postulates of thermodynamics that the knowledge of the entropy S of a specific homogeneous system allows for the prediction of its thermodynamic behavior, provided S is expressed as a function of the *extensive* quantities $(X = E_0, X_1, \ldots, X_r)$ that define its thermodynamic state. The equation

$$S = S(X_0 = E, X_1, \ldots, X_r) \tag{2.54}$$

is called the **fundamental equation**, and represents a *complete* description of the thermodynamics of the system being considered.

From the fundamental equation, one can obtain the expression of the temperature T and of the generalized forces f by taking derivatives with respect to the extensive quantities. These expressions are called, generically, **equations of state**.

For example, if we take the derivative of the fundamental equation with respect to the internal energy E, we obtain the temperature T expressed as a function of the internal energy E, the number of particles N, and the volume V. Then, if we take the derivative with respect to V, we obtain p/T expressed as a function of the same variables. We are thus able evaluate the pressure p. One can then express E as a function of T and of the other variables from the first relation and substitute it in the second one to obtain the habitual equation of state:

$$p = p(T, V, N). \tag{2.55}$$

Knowledge of this equation of state does not allow us to reconstruct the fundamental equation. Indeed, in order to exploit the relation (for $N = \text{const.}$)

$$\mathrm{d}S = \frac{\partial S}{\partial E}\Big)_{V,N} \mathrm{d}E + \frac{\partial S}{\partial V}\Big)_{E,N} \mathrm{d}V = \frac{p}{T}\,\mathrm{d}E + \frac{1}{T}\,\mathrm{d}V, \tag{2.56}$$

to obtain S, we need to express E as a function of T, V, and N. Now, setting $\mathrm{d}S = 0$ in this equation, we have

$$\frac{\partial E}{\partial V}\Big)_{S,N} = \frac{\partial S/\partial V)_{E,N}}{\partial S/\partial E)_{V,N}} = -T \cdot \frac{p}{T} = -p. \tag{2.57}$$

However, this is not enough allow the evaluation of $\partial E/\partial V)_{T,N}$. One has in fact

$$\begin{aligned} \frac{\partial E}{\partial V}\Big)_{T,N} &= \frac{\partial E}{\partial V}\Big)_{S,N} + \frac{\partial E}{\partial S}\Big)_{V,N} + \frac{\partial E}{\partial S}\Big)_{V,N} \frac{\partial S}{\partial V}\Big)_{T,N} \\ &= -p + T \cdot \frac{\partial S}{\partial V}\Big)_{T,N}, \end{aligned} \tag{2.58}$$

where the last term cannot be evaluated from the given equation of state alone.

Exercise 2.2 The following expressions are proposed as possible fundamental equations for various thermodynamic systems. However, some of them violate some of the postulates of entropy and are therefore not acceptable from a physics standpoint. Find the five that are not acceptable from a physics standpoint, and indicate which postulates are violated by each one. The quantities κ and θ are positive constants. When fractional exponents appear, only the positive values of the roots should be considered.

1. $S = \kappa\,(NVE)^{1/3}$.
2. $S = \kappa[NE/V]^{2/3}$.
3. $S = \kappa\left[NE - \theta V^2\right]^{1/2}$.
4. $S = \kappa\left[V^3/NE\right]$.
5. $S = \kappa\left[N^2 VE^2\right]^{1/5}$.
6. $S = N\kappa \ln\left(VE/\theta N^2\right)$.
7. $S = \kappa\,[NE]^{1/2} \exp\left(-V^2/(\theta N^2)\right)$.
8. $S = \kappa\,[NE]^{1/2} \exp\left(-EV/(\theta N)\right)$.

Exercise 2.3 Let the fundamental equation for systems A and B be given by

$$S = \kappa\,(NVE)^{1/3}.$$

Let $N^{(\mathrm{A})}$ and $V^{(\mathrm{A})}$ be the number of particles and volume of system A, and similarly for B. Assume that the total internal energy of the system $\mathrm{A} \cup \mathrm{B}$ is equal to E_0, and that the

systems can exchange energy but neither particles nor volume. Evaluate the internal energies of the two systems at equilibrium.

2.9 Energy Scheme

The fundamental equation we discussed is the most natural, even from the point of view of statistical mechanics. However, essentially for historical reasons, we prefer to use a different (but equivalent) formulation of the fundamental principle of thermodynamics, in which entropy assumes the role of an independent variable, while energy becomes a dependent variable that satisfies a variational principle. This formalism is known as the **energy scheme**.

In this formalism, the maximum entropy principle—

> *Among all states with a specific value of* internal energy*, the state of equilibrium is that in which* entropy is maximal

— is replaced by the principle of minimum internal energy—

> *Among all states with a specific* entropy *value, the state of equilibrium is that in which* internal energy *is minimal.*

For the proof, it can be useful to refer to figure 2.3. Let ΔX be a virtual variation of the extensive variables (excluding the internal energy E) with respect to the equilibrium value X^{eq}. Then,

$$\Delta S = S(E, X^{\text{eq}} + \Delta X) - S(E, X^{\text{eq}}) \leq 0, \tag{2.59}$$

because of the maximum entropy principle. Since S is a monotonically increasing function of E, there exists a value $E' \geq E$ such that $S(E', X^{\text{eq}} + \Delta X) = S(E, X^{\text{eq}})$. Therefore, if S is kept constant, as the system moves out of equilibrium, E cannot but increase. This is what we wanted to prove.

Therefore, the fundamental equation in the energy scheme is defined when the dependence of the internal energy E on the entropy S and on the other extensive variables $(X_1, \ldots, X_r)$ is known:

$$E = E(S, X_1, \ldots, X_r). \tag{2.60}$$

The differential of E assumes the form

$$\mathrm{d}E = T\mathrm{d}S + \sum_{i=1}^{r} f_i \mathrm{d}X_i. \tag{2.61}$$

Comparing this equation with eq. (2.34) we can identify $-\sum_i f_i \mathrm{d}X_i$ with the infinitesimal work $đW$ that the system performs during an infinitesimal transformation. Then $T\mathrm{d}S$ is identified with the infinitesimal heat $đQ$ released to the system in this transformation.

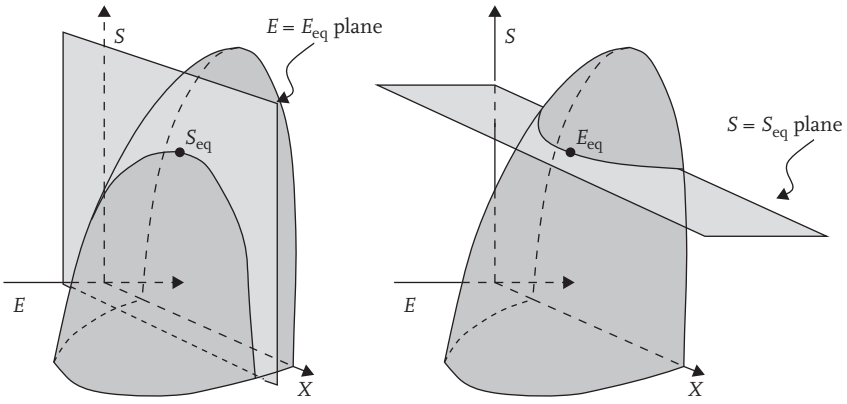


FIGURE 2.3. A schematic representation of the fundamental surface. The equilibrium state can be characterized by the maximum entropy on the plane $E = \text{const.}$ (left) or, equivalently, by the minimum internal energy on the plane $S = \text{const.}$ (right).

We can thus recognize in this expression the differential version of the first principle of thermodynamics.

Whereas the entropy S is a concave function of the extensive variables $\{X_i\}$, the internal energy E is a *convex* function of S and of the other extensive variables X_i, $i = 1, \ldots, r$. We know in fact that if $X^{(1)} = (X_0^{(1)} = E^{(1)}, X_1, \ldots, X_r^{(1)})$ and $X^{(2)} = (X_0^{(2)} = E^{(2)}, X_1^{(2)}, \ldots, X_r^{(2)})$ are two states of equilibrium, and $0 \leq \alpha \leq 1$, one has

$$S((1-\alpha)X^{(1)} + \alpha X^{(2)}) \leq (1-\alpha)\, S(X^{(1)}) + \alpha\, S(X^{(2)}), \tag{2.62}$$

which expresses the concavity of entropy. Let us denote by $S^{(1)}$ the value of $S(X^{(1)})$ and by $S^{(2)}$ the value of $S(X^{(2)})$. We can then define the states of equilibrium in the energy scheme as follows:

$$Y^{(1)} = \left(S^{(1)}, X_1^{(1)}, \ldots, X_r^{(1)}\right); \qquad Y^{(2)} = \left(S^{(2)}, X_1^{(2)}, \ldots, X_r^{(2)}\right). \tag{2.63}$$

We get therefore

$$S\left((1-\alpha)E^{(1)} + \alpha E^{(2)}, \ldots, (1-\alpha)X_r^{(1)} + \alpha X_r^{(2)}\right) \geq (1-\alpha)S^{(1)} + \alpha S^{(2)}. \tag{2.64}$$

Define E' such that

$$S(E', (1-\alpha)X_1^{(1)} + \alpha X_1^{(2)}, \ldots, (1-\alpha)X_r^{(1)} + \alpha X_r^{(2)}) = (1-\alpha)S^{(1)} + \alpha S^{(2)}. \tag{2.65}$$

Then, since S is an increasing function of E, we have

$$E' \leq (1-\alpha)E^{(1)} + \alpha E^{(2)}. \tag{2.66}$$

Therefore

$$E' \leq E\left((1-\alpha)Y^{(1)} + \alpha Y^{(2)}\right) \leq (1-\alpha)E(Y^{(1)}) + \alpha E(Y^{(2)}), \tag{2.67}$$

which implies that the internal energy function is convex.

Exercise 2.4 For each of the acceptable fundamental equations in exercise 2.2, express E as a function of S, V, and N.

2.10 Intensive Variables and Thermodynamic Potentials

The derivatives $f_i = \partial E/\partial X_i$ of the internal energy E with respect to extensive quantities S and $\{X_i\}$ are called **intensive quantities** (or **fields**). For uniformity, we define $f_0 = T = \partial E/\partial S$. A given quantity f_i is called the **conjugate** of the corresponding variable X_i, and vice versa. The temperature T and entropy S are therefore conjugates, as are the pressure, with the *opposite sign* ($-p$), and the volume V.

Since both E and X_i are extensive, in a homogeneous system, intensive variables do not depend on system size. Moreover, if a system is composed of several subsystems that can exchange the extensive quantity X_i, the corresponding intensive variable f_i assumes at equilibrium the same value in the subsystems that are in contact.

One needs to distinguish between intensive variables, which possess this property, and **densities**, which are quantities independent of system size and which are obtained by taking the ratio of two extensive variables in a homogenous system. For instance, if N is the number of particles, the quantities $s = S/N$, $v = V/N$, and $\epsilon = E/N$ are respectively the entropy, volume, and internal energy densities per particle, and do not depend on the system size. Analogously, one can define densities with respect to volume, such as $s_V = S/V$. Unlike intensive variables, the densities can assume different values in systems that mutually in contact.

We now want to identify the state of equilibrium among all states that exhibit a given value of an intensive variable f_i. Specifically, for $i = 0$, we are confronted with the case of a system with a fixed temperature—in practice this can be realized by putting the system in contact with a much larger system (a reservoir) via a diathermal wall. In these conditions, the reservoir's temperature does not, practically speaking, depend on the quantity of energy it exchanges with the system, and so the latter, in equilibrium, assumes the temperature value established by the reservoir.

Let us define the **Helmholtz free energy** $F(T, X)$ by the relation

$$F(T, X) = E(S(T, X), X) - TS(T, X), \tag{2.68}$$

where $X = (X_1, \ldots, X_r)$ is the list of the extensive thermodynamic variables, apart from the energy, that identify the equilibrium state, and $S(T, X)$ is the value of the entropy in the equilibrium state identified by T and X. Since $T = \partial E/\partial S)_X$, $F(T, X)$ is the *opposite* of the **Legendre transform** of the internal energy $E(S, X)$ with respect to S. (The properties of the Legendre transformation are summarized in secs. B.3 and B.4.) We can then show that the thermodynamical equilibrium in these conditions is characterized by the following variational principle:

The value of the Helmholtz free energy is minimal for the equilibrium state among all virtual states at the given temperature T.

To prove this result, let us denote by X the extensive variables that are free to vary among different virtual states, and let us evaluate the free energy difference between the equilibrium state identified by (T, X^{eq}) and a generic state identified by (T, X). (We do not consider explicitly further extensive variables that are kept fixed.) We obtain

$$\Delta F = E(S(T,X),X) - E(S(T,X^{eq}),X^{eq}) - T\,[S(T,X) - S(T,X^{eq})]. \tag{2.69}$$

Let us add and subtract $E(S(T,X^{eq}),X)$ on the right-hand side. We obtain

$$\begin{aligned}\Delta F = {} & [E(S(T,X),X) - E(S(T,X^{eq}),X)] \\ & + [E(S(T,X^{eq}),X) - E(S(T,X^{eq}),X^{eq})] \\ & - T\,[S(T,X) - S(T,X^{eq})].\end{aligned} \tag{2.70}$$

The second term within brackets is nonnegative, because of the variational principle satisfied by the internal energy. Let us consider the first and last terms. Taking into account that E is convex and that $T = \partial E/\partial S)_X$, we obtain

$$E(S(T,X),X) - E(S(T,X^{eq}),X) \geq T[S(T,X) - S(T,X^{eq})]. \tag{2.71}$$

We have, therefore,

$$\Delta F \geq 0, \tag{2.72}$$

as we intended to prove. Therefore, knowledge of F as a function of the temperature T and of the extensive variables X (excluding the entropy) allows us to solve the fundamental problem of thermodynamics. In other words $F = F(T, X_1, \ldots, X_r)$ is an equivalent form of the fundamental equation.

We can generalize the argument to other intensive variables as follows. Consider, for example, the function $\Phi(S, f_1, X_2, \ldots, X_r)$ defined by

$$\begin{aligned}\Phi(S, f_1, X_2, \ldots, X_r) = {} & E(S, X_1(S, f_1, X_2, \ldots, X_r), X_2, \ldots, X_r) \\ & - f_1 X_1(S, f_1, X_2, \ldots, X_r),\end{aligned} \tag{2.73}$$

where $X_1(S, f_1, X_2, \ldots, X_r)$ is determined by the condition

$$f_1 = \left.\frac{\partial E}{\partial X_1}\right|_{X_1 = X_1(S, f_1, X_2, \ldots, X_r)}. \tag{2.74}$$

The function $\Phi(S, f_1, X_2, \ldots, X_r)$ is the opposite of the Legendre transform of E with respect to X_1. We then have the following criterion for the equilibrium state:

Among all the states that have the same value of f_1, the equilibrium state corresponds to the minimum of Φ.

Because of this criterion, the function Φ is known as the **thermodynamic potential**, since it reaches the minimum value in thermodynamic equilibrium, analogously to mechanical potential, which reaches its minimum in mechanical equilibrium.

This result can be proved by following the same path as above. Let us suppose that we set the value of f_1, and we allow some other extensive variables X to vary. We then get

$$\Phi(S, f_1, X) = E(S, X_1(f_1, X), X) - f_1 X_1(S, f_1, X). \tag{2.75}$$

Denote the equilibrium values of X_1 by X_1^{eq} and of X by X^{eq}. To a virtual variation ΔX of X there corresponds an analogous variation $\Delta X_1 = X_1(S, f_1, X) - X_1^{\mathrm{eq}}$ of X_1. We then have

$$\begin{aligned}
\Phi(S, f_1, X^{\mathrm{eq}}) &= E(S, X_{1\mathrm{eq}+\Delta X_1}, X^{\mathrm{eq}} + \Delta X) - f_1 \left(X_1^{\mathrm{eq}} + \Delta X_1\right) \\
&\geq E(S, X_1^{\mathrm{eq}}, X^{\mathrm{eq}} + \Delta X) + \left.\frac{\partial E}{\partial X_1}\right|_{X_1^{\mathrm{eq}}, X^{\mathrm{eq}}+\Delta X} \\
&\qquad - f_1 \left(X_1^{\mathrm{eq}} + \Delta X_1\right) \\
&= E(S, X_1^{\mathrm{eq}}, X^{\mathrm{eq}} + \Delta X) - f_1 X_1^{\mathrm{eq}} \\
&\geq E(S, X_1^{\mathrm{eq}}, X^{\mathrm{eq}}) - f_1 X_1^{\mathrm{eq}} \\
&= \Phi(S, f_1, X^{\mathrm{eq}}).
\end{aligned} \tag{2.76}$$

In this derivation, we applied in the second line the convexity of E, and in the fifth line the variational principle in the energy scheme.

Let us observe that, by eq. (B.22), the partial derivative of Φ, performed with respect to f_1, with the other extensive variables X kept fixed, yields the value of the extensive variable X_1 with the opposite sign:

$$\begin{aligned}
\left.\frac{\partial \Phi}{\partial f_1}\right)_{S,X} &= \left.\frac{\partial E}{\partial X_1}\right)_{S,X} \left.\frac{\partial X_1}{\partial f_1}\right)_{S,X} - X_1(f_1, X) - f_1 \left.\frac{\partial X_1}{\partial f_1}\right)_{S,X} \\
&= -X_1(f_1, X).
\end{aligned} \tag{2.77}$$

Nothing prevents us from considering two or more intensive variables as fixed—for example, $f_0 = T$ and f_1. Similar considerations will then lead us to introduce the thermodynamic potential $\Phi(T, f_1, X)$, related to the Legendre transform of E with respect to both S and X_1:

$$\begin{aligned}
\Phi(T, f_1, X) = E(&S(T, f_1, X), X_1(T, f_1, X), X) \\
&- TS(T, f_1, X) - f_1 X_1(T, f_1, X).
\end{aligned} \tag{2.78}$$

This thermodynamic potential reaches its minimum value at the equilibrium state, among all the states with the same values of T and f_1. We can therefore obtain a whole set of thermodynamic potentials, by performing a Legendre transform with respect to some of the extensive variables X_i, $i = 1, \ldots, r$. We cannot, however, eliminate *all* extensive variables in this manner. We see later that if we did this, the resulting thermodynamic potential would identically vanish. For the time being, it is sufficient to observe that the Φ potentials are extensive quantities, and one cannot see how they could be a function only of intensive quantities like the fs.

For each thermodynamic potential, there is a set of "natural" variables—some are *intensive*, that is, the conjugates of the variables with respect to which the Legendre

transformation is performed, and the others are *extensive*. One can readily see that the thermodynamic potential

$$\Phi(f_0 = T, f_1, f_2, \ldots, f_k, X_{k+1}, \ldots, X_r) = E - \sum_{i=0}^{k} f_i X_i, \tag{2.79}$$

where the extensive variables $(X_0 = S, X_1, \ldots, X_k)$ have been swapped for their conjugate variables $(f_0, f_1, \ldots, f_k)$, is *convex* as a function of the remaining extensive variables $(X_{k+1}, \ldots, X_r)$ for fixed values of the intensive variables $(f_1, \ldots, f_k)$. More precisely, for any small virtual variation $\Delta X = (\delta X_{k+1}, \ldots, \delta X_r)$ one has

$$\delta^2 \Phi = \sum_{i,j=k+1}^{r} \frac{\partial^2 \Phi}{\partial X_i \partial X_j}\Bigg)_{f_0,\ldots,f_k} \delta X_i \delta X_j \geq 0, \tag{2.80}$$

that is, the matrix of second derivatives $(\partial^2 \Phi / \partial X_i \partial X_j)$ is **positive semidefinite**. On the other hand, Φ is a *concave* function of the intensive variables f when the extensive variables X are fixed. This follows from the properties of the Legendre transformation (see sec. B.4) and its proof is left as an exercise.

Exercise 2.5 Consider the thermodynamic potential

$$\Phi(f_0 = T, f_1, f_2, \ldots, f_k, X_{k+1}, \ldots, X_r) = E - \sum_{i=0}^{k} f_i X_i.$$

Prove that the matrix of the second derivatives $(\partial^2 \Phi / \partial f_i \partial f_j)$ is *negative semidefinite* at each point.

2.11 Free Energy and Maxwell Relations

We can now be more explicit. If we take the temperature T as an independent variable in place of the entropy S, we obtain, as the relevant thermodynamic potential, the Helmholtz free energy, which we have defined in eq. (2.68). In this equation, the equilibrium value $S(T, X)$ of the entropy in the equilibrium state identified by (T, X) is implicitly defined by the relation

$$T = \frac{\partial E}{\partial S}\bigg|_{S=S(T,X)}. \tag{2.81}$$

As we have seen, the Helmholtz free energy takes on its minimal value for the equilibrium state, among all virtual states corresponding to the same value of the temperature T. The natural variables of F are the temperature T and the extensive variables $X_1, \ldots, X_r$, entropy excluded.

The entropy is obtained by taking the derivative of F with respect to T:

$$\frac{\partial F}{\partial T}\Bigg)_{X_1,\ldots,X_r} = -S. \tag{2.82}$$

We thus obtain the expression for the differential of F:

$$\mathrm{d}F = -S\mathrm{d}T + \sum_{i=1}^{r} f_i \mathrm{d}X_i. \tag{2.83}$$

More specifically, by setting $X_1 = V$, one has

$$-p = \frac{\partial F}{\partial V}\Big)_{V,X_2,\ldots,X_r}, \tag{2.84}$$

where $X_2, \ldots, X_r$ are the other extensive variables (excluding both entropy and volume).

If we now take the derivative of this equation with respect to T and we use the theorem of the equality of mixed derivatives, we obtain

$$\frac{\partial p}{\partial T}\Big)_{V,X_2,\ldots,X_r} = -\frac{\partial^2 F}{\partial T \partial V} = -\frac{\partial^2 F}{\partial V \partial T} = \frac{\partial S}{\partial V}\Big)_{T,X_2,\ldots,X_r}. \tag{2.85}$$

This is a thermodynamic identity that relates the heat absorbed by a system in isothermal expansion with the variation in pressure and temperature at constant volume. In effect, if in an isothermal expansion volume varies by $\mathrm{d}V$, the heat absorbed is given by $đQ = T\,\partial S/\partial V)_T \mathrm{d}V$. We thus obtain

$$\frac{1}{T}\,\frac{đQ}{\mathrm{d}V}\Big)_T = \frac{\partial p}{\partial T}\Big)_V. \tag{2.86}$$

These relations between thermodynamic derivatives that follow from the equality of mixed derivatives of thermodynamic potentials are called **Maxwell relations**.

The free energy designation is derived from the following property. If a system is put in contact with a reservoir at temperature T, the maximum amount of work $W_{\max}$ that it can perform on its environment is equal to the variation in free energy between the initial and final states. In other words, one has

$$W \leq W_{\max} = F_{\mathrm{in}} - F_{\mathrm{fin}}. \tag{2.87}$$

More specifically, if the system relaxes toward equilibrium, $F_{\mathrm{fin}} = F_{\mathrm{eq}} \leq F_{\mathrm{in}}$, and therefore, $W_{\max} \geq 0$—in other words, the system is able to perform work on its environment.

In order to prove this assertion, let us suppose that the global system (the actual system plus the reservoir) is adiabatically isolated. In this situation, we have

$$W = E_{\mathrm{in}} - E_{\mathrm{fin}}, \tag{2.88}$$

where E_{fin} (E_{in}) is the final (initial) internal energy of the global system. On the other hand, one has

$$S_{\mathrm{fin}} \geq S_{\mathrm{in}}, \tag{2.89}$$

where S_{fin} (S_{in}) is the final (initial) entropy of the global system. Expressed in greater detail, we have

$$S_{\mathrm{fin}} = S_{\mathrm{fin}}^{(\mathrm{s})} + S_{\mathrm{fin}}^{(\mathrm{r})} \geq S_{\mathrm{in}}^{(\mathrm{s})} + S_{\mathrm{in}}^{(\mathrm{r})}, \tag{2.90}$$

where (s) and (r) refer to system and reservoir, respectively. However, since (r) is a heat reservoir at the fixed temperature T, one has

$$S^{(r)}_{\text{fin}} = S^{(r)}_{\text{in}} + \frac{\partial S^{(r)}}{\partial E^{(r)}} \left(E^{(r)}_{\text{fin}} - E^{(r)}_{\text{in}} \right) = S^{(r)}_{\text{in}} + \frac{1}{T} \left(E^{(r)}_{\text{fin}} - E^{(r)}_{\text{in}} \right). \tag{2.91}$$

From eq. (2.88) we obtain

$$E^{(r)}_{\text{fin}} - E^{(r)}_{\text{in}} = W - E^{(s)}_{\text{fin}} + E^{(s)}_{\text{in}}, \tag{2.92}$$

and therefore

$$S^{(r)}_{\text{fin}} - S^{(r)}_{\text{in}} = -\frac{1}{T} \left(W - E^{(s)}_{\text{fin}} + E^{(s)}_{\text{in}} \right) \geq S^{(s)}_{\text{in}} - S^{(s)}_{\text{fin}}, \tag{2.93}$$

and, since $T \geq 0$, multiplying by $-T$, we obtain

$$W \leq \left(E^{(s)}_{\text{in}} - TS^{(s)}_{\text{in}} \right) - \left(E^{(s)}_{\text{fin}} - TS^{(s)}_{\text{fin}} \right) = F_{\text{in}} - F_{\text{fin}}. \tag{2.94}$$

This is what we wanted to prove.

2.12 Gibbs Free Energy and Enthalpy

The thermodynamic potential that depends on the temperature T and the pressure p and on the other extensive variables X (except the entropy S and the volume V) is the **Gibbs free energy**. It is given by

$$G(T, p, X) = E - TS + pV = F + pV. \tag{2.95}$$

It is related to the Legendre transform of F with respect to V. One has to keep in mind that the intensive variable conjugate to V is the opposite of pressure, that is, $(-p)$. This explains the sign before pV in this expression.

The Gibbs free energy satisfies the following variational principle:

Among all states that have the same temperature *and* pressure *values, the state of equilibrium is that in which the Gibbs free energy assumes the minimum value.*

This principle can be proved by following the lines of the variational principle of F. The proof is left as an exercise.

Exercise 2.6 Prove the variational principle for the Gibbs free energy.

One can also see, moreover, that

$$\frac{\partial G}{\partial T}\bigg)_{p,X} = -S, \qquad \frac{\partial G}{\partial p}\bigg)_{T,X} = V. \tag{2.96}$$

Therefore, the differential of G is expressed by

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}p + \sum_{i=2}^{r} f_i \mathrm{d}X_i. \tag{2.97}$$

It is also easy to prove that if a system is brought toward equilibrium while temperature and pressure are kept constant, the maximum work that can be performed on its environment is given by the difference between the initial and final values of G. This is what justifies its designation as *free energy*.

Exercise 2.7 Prove the preceding assertion.

We can also swap V for p in the internal energy E by a Legendre transformation. We thus obtain a new thermodynamic potential called **enthalpy**, usually denoted by H:

$$H(S, p, X) = E + pV. \tag{2.98}$$

Enthalpy governs the equilibrium of *adiabatic* processes occurring at constant pressure:

> *Among all states that have the same* entropy *and* pressure *values, the state of equilibrium is the one that corresponds to the minimum value of enthalpy.*

Enthalpy is particularly important for chemical reactions in open containers (which are therefore subject to atmospheric pressure).

Exercise 2.8 Prove that the heat Q released by a system that is relaxing at equilibrium at constant pressure is less than (or equal to) its change in enthalpy:

$$Q \leq H_{\mathrm{in}} - H_{\mathrm{fin}} = -Q_{\mathrm{rev}}.$$

If the system (s) is in contact with a reservoir (r) at pressure p, the heat released by the system to the reservoir in a transformation is equal to the opposite of the change in its enthalpy. Note that we are *not* supposing that the system is in thermal equilibrium with the reservoir, only that the reservoir exchanges energy with the system via either heat or its volume change. Indeed, the heat received by the reservoir is given by

$$\begin{aligned} Q &= \Delta E^{(\mathrm{r})} + W^{(\mathrm{r})} \\ &= \Delta E^{(\mathrm{r})} + p\,\Delta V^{(\mathrm{r})} = -\left(\Delta E^{(\mathrm{s})} + p\,\Delta V^{(\mathrm{s})}\right), \end{aligned} \tag{2.99}$$

since $\Delta E^{(\mathrm{r})} = -\Delta E^{(\mathrm{s})}$ and $\Delta V^{(\mathrm{r})} = -\Delta V^{(\mathrm{s})}$. When p is constant, we have

$$\Delta H^{(\mathrm{s})} = \Delta E^{(\mathrm{s})} + \Delta\left(pV^{(\mathrm{s})}\right) = \Delta E^{(\mathrm{s})} + p\,\Delta V^{(\mathrm{s})}. \tag{2.100}$$

Comparing this equation with that above, we obtain the desired result.

The differential of the enthalpy is given by

$$\mathrm{d}H = T\mathrm{d}S + V\mathrm{d}p + \sum_{i=2}^{r} f_i \mathrm{d}X_i. \tag{2.101}$$

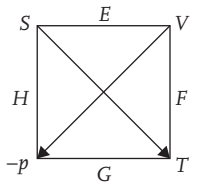


FIGURE 2.4. The Koenig-Born diagram for the variables entropy S and volume V. The thermodynamic potentials on the sides are accompanied by their respective "natural" variables at the vertices. Note that the potentials follow one another in alphabetical order, in a clockwise pattern, and that the conjugate variable of volume V is given by $-p$, following our conventions. The arrows go from the extensive variables to their conjugates.

The equality of the mixed derivatives of G and H yields two more Maxwell relations:

$$\left.\frac{\partial S}{\partial p}\right)_T = -\left.\frac{\partial V}{\partial T}\right)_p, \qquad \left.\frac{\partial T}{\partial p}\right)_S = \left.\frac{\partial V}{\partial S}\right)_p. \tag{2.102}$$

2.13 The Koenig-Born Diagram

Relations between thermodynamic potentials, extensive variables, and the corresponding intensive variables are summarized in the **Koenig-Born diagram**, which is shown in figure 2.4. In this diagram, the thermodynamic potentials are written on the sides of a square, while the relevant thermodynamic variables are written on the corners, in such a fashion that each thermodynamic potential is accompanied by its natural variables. The diagonals associate conjugated pairs of variables like (S with T, and V with $(-p)$), and the arrows determine the sign of the corresponding differential element—if the arrow points away from the potential being considered, the sign is $(+)$; otherwise, it is $(-)$. So the differential of E is $T\mathrm{d}S + (-p)\mathrm{d}V$, while the differential of H is $T\mathrm{d}S - V\mathrm{d}(-p)$.

It is easy to read Maxwell's relations from this diagram: as shown, for example, in fig. 2.5, they correspond to triangles that share one side—for example, the relation

$$\left.\frac{\partial S}{\partial(-p)}\right)_T = \left.\frac{\partial V}{\partial T}\right)_p \tag{2.103}$$

corresponds to the bottom side, which can be obtained from the equality of the second-order derivatives of the Gibbs free energy G. There are no extra signs to account for, since the arrows' structure is the same for both derivatives. If we instead consider the F side, we obtain

$$\left.\frac{\partial(-p)}{\partial T}\right)_V = -\left.\frac{\partial S}{\partial V}\right)_T. \tag{2.104}$$

Here there is an explicit "minus" sign, because on the one hand we have one arrow that points *toward* the side being considered, and on the other, one that points *away* from it.

More formally, the structure of a Maxwell relation is the following. Let X and Y be two extensive variables, whose conjugate intensive variables are f and g, respectively. We then have, for example, $\partial X/\partial g)_f = \partial Y/\partial f)_g$ and $\partial X/\partial Y)f = \partial g/\partial f)_Y$. The sign change in this last relation comes about because a derivative with respect to an extensive variable is swapped with a derivative with respect to an intensive one. Note that the quantity kept

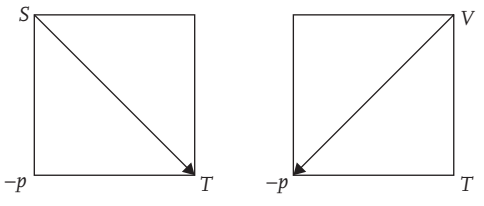


FIGURE 2.5. Koenig-Born diagram and Maxwell relations. The two schemes represent the relation $\partial S/\partial(-p))_T = \partial V/\partial T)_p$.

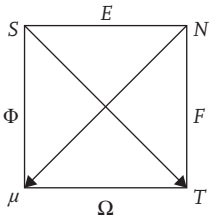


FIGURE 2.6. Koenig-Born diagram relative to the variables S and N.

constant is always the conjugate of the quantity whose derivative is taken. We always take $-p$ to be the conjugate variable to V.

2.14 Other Thermodynamic Potentials

The Legendre transforms with respect to the number of particles N (for each chemical species) produce other thermodynamic potentials, which are extremely important in chemistry and statistical mechanics but which do not have either a name or a standard notation.

The Legendre transform of F with respect to N produces a thermodynamic potential (often written as Ω) that depends on T, on the volume V, on the chemical potential μ, and on the other extensive variables:

$$\Omega(T, V, \mu) = F(T, V, N) - \mu N. \tag{2.105}$$

Its differential is given by

$$\mathrm{d}\Omega = -S\mathrm{d}T - p\mathrm{d}V - N\mathrm{d}\mu. \tag{2.106}$$

If one instead transforms E, one obtains a rarely used potential that depends on S, V, and μ, which we will denote by Φ:

$$\Phi(S, V, \mu) = E(S, V, N) - \mu N. \tag{2.107}$$

Its differential is given by

$$\mathrm{d}\Phi = T\mathrm{d}S - p\mathrm{d}V - N\mathrm{d}\mu. \tag{2.108}$$

The E, F, Ω, and Φ potentials allow one to write a Koenig-Born diagram analogous to the one we saw in the preceding section, as shown in fig. 2.6. We therefore arrive, for

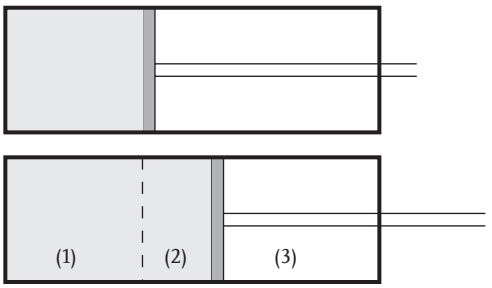


FIGURE 2.7. The measurement of the chemical potential. The solute is kept in the region left of the semipermeable piston, and exerts an osmotic pressure toward the right. The piston is then very gently moved to the right, allowing the solute to occupy region (2). The whole system is enclosed in adiabatic walls.

example, at the following Maxwell relations:

$$\left.\frac{\partial S}{\partial \mu}\right)_T = \left.\frac{\partial N}{\partial T}\right)_\mu , \tag{2.109}$$

$$\left.\frac{\partial S}{\partial N}\right)_T = \left.\frac{\partial \mu}{\partial T}\right)_N . \tag{2.110}$$

2.15 The Measurement of the Chemical Potential

We now show how, by using a suitable thermodynamic potential, one can relate the chemical potential difference to the osmotic pressure, thus allowing for its mechanical measurement.

Let us consider a cylinder, inside of which there is a piston fitted with a semipermeable membrane that blocks the passage of chemical species A but allows the passage of the solvent (see figure 2.7). The system is adiabatically isolated from the external environment. The system exerts an osmotic pressure p_A on the piston, which we regard as positive when it acts toward the right.

Let us now suppose that the piston is shifted reversibly by a small interval toward the right. Let us mentally subdivide the system into three subsystems, from left to right:

1. The system that occupies the area to the left of the initial position of the piston.
2. The system that occupies the area between the two positions of the piston.
3. The system that occupies the area to the right of the final position of the piston.

Since the transformation is reversible and the system is adiabatically isolated, the total variation of entropy ΔS vanishes, and the variation of the internal energy ΔE is given by

$$\Delta E = -p_A V^{(2)}. \tag{2.111}$$

Let us now consider the thermodynamic potential (which we denote by $\mathcal{F}$) whose independent variables are the volume V, the temperature T, the chemical potential of the solvent μ_s, and the number N of solute particles. It is defined by

$$\mathcal{F}(T, V, \mu_s, N) = E - TS - \mu_s N_s, \tag{2.112}$$

where N_s is the number of solvent particles.

The chemical potential of the solute is given by

$$\mu = \frac{\partial \mathcal{F}}{\partial N}\Big)_{T,V,\mu_s} = \mu(T, \mu_s, \rho_A), \tag{2.113}$$

where $\rho_A = N/V$ is the concentration of the solute. It is clear that μ depends on N only via ρ_A, since it is an intensive quantity. Since the system is at equilibrium at the beginning and the end of the transformation, and the transformation takes place at constant entropy S and N_s, we have

$$\Delta\mathcal{F} = \Delta\mathcal{F}^{(1)} + \Delta\mathcal{F}^{(2)} + \Delta\mathcal{F}^{(3)} = \Delta E = -p_A V^{(2)}. \tag{2.114}$$

In addition, one has $T^{(1)} = T^{(2)} = T^{(3)} = \text{const.}$, while the chemical potential μ_s of the solvent is the same in the three systems we are considering. It is obvious that $\Delta\mathcal{F}^{(3)} = 0$. Therefore, the variation in the total free energy $\mathcal{F}$ is due to the shift of ΔN particles from subsystem (1) to subsystem (2), where

$$\Delta N = (\rho_A^{(1)} - \rho_A^{(3)}) V^{(2)} = \Delta\rho_A V^{(2)}. \tag{2.115}$$

Here $\rho_A^{(i)}$ is the solute concentration in system (i) $(i = 1, 3)$. System (1) undergoes a variation

$$\Delta\mathcal{F}^{(1)} = -\mu^{(1)}\,\Delta N = -\mu^{(1)}(\rho_A^{(1)} - \rho_A^{(3)}) V^{(2)} = -\mu^{(1)}\Delta\rho_A V^{(2)}. \tag{2.116}$$

On the other hand, in system (2) the concentration changes from $\rho_A^{(3)}$ to $\rho_A^{(1)}$. We thus have

$$\begin{aligned}\Delta\mathcal{F}^{(2)} &= V^{(2)} \int_{\rho_A^{(3)}}^{\rho_A^{(1)}} \mathrm{d}\rho\, \mu(T, \mu_s, \rho) \\ &\simeq \int_{\rho_A^{(3)}}^{rho_A^{(1)}} \mathrm{d}\rho \left[\mu^{(3)} + \left(\mu^{(1)} - \mu^{(3)}\right) \frac{\rho - \rho_A^{(3)}}{\Delta\rho_A}\right] \\ &= \frac{1}{2} V^{(2)} \Delta\rho_A \left(\mu^{(1)} + \mu^{(3)}\right),\end{aligned} \tag{2.117}$$

where we have used a linear approximation for the dependence of μ on ρ. We thus obtain

$$\begin{aligned}\Delta\mathcal{F} &= \Delta\mathcal{F}^{(1)} + \Delta\mathcal{F}^{(2)} \simeq -\mu^{(1)}\Delta\rho_A V^{(2)} + \frac{1}{2} V^{(2)} \Delta\rho_A \left(\mu^{(1)} + \mu^{(3)}\right) \\ &= \frac{1}{2}\Delta\rho_A \left(\mu^{(3)} - \mu^{(1)}\right).\end{aligned} \tag{2.118}$$

Comparing with eq. (2.114), we obtain

$$p_A \simeq \frac{1}{2} \left(\mu^{(1)} - \mu^{(3)}\right) \Delta\rho_A. \tag{2.119}$$

Therefore, if the concentration difference is known, the measurement of the osmotic pressure yields the difference in the chemical potentials.

2.16 The Euler and Gibbs-Duhem Equations

Let us imagine that we are putting a homogeneous system, characterized by the extensive variables $X_0 = S, X_1, \ldots, X_r$ and the corresponding internal energy value E, in contact with other $(m-1)$ identical systems. Due to the extensivity of the variables X and the function E, we obtain a system characterized by the values $X_i' = mX_i$ of the extensive variables and by the internal energy value $E' = mE$. Following the same line of reasoning, we can mentally subdivide the system into n identical systems, each with the values $(1/n)X_i$ for the X_i variables and with the value $(1/n)E$ for the internal energy. Because of continuity, we can conclude that internal energy E, since it is an extensive function of the extensive variables X_i, must, for each positive factor λ, satisfy a homogeneity relation:

$$E(\lambda S, \lambda X_1, \ldots, \lambda X_r) = \lambda E(S, X_1, \ldots, X_r). \tag{2.120}$$

By taking the derivative of this equation with respect to λ and setting $\lambda = 1$, we obtain the **Euler equation**:

$$S\frac{\partial E}{\partial S} + \sum_{i=1}^{r} X_i \frac{\partial E}{\partial X_i} = E. \tag{2.121}$$

Making it explicit, we obtain

$$TS + \sum_{i=1}^{r} f_i X_i = E. \tag{2.122}$$

It follows from this equation that the Legendre transform of E with respect to *all* extensive variables vanishes identically.

More particularly, for *simple fluids*, one obtains

$$TS - pV + \mu N = E, \tag{2.123}$$

which among other things implies that

$$\mu = (E - TS + pV)/N = G/N, \tag{2.124}$$

$$\Omega = E - TS - \mu N = -pV. \tag{2.125}$$

We emphasize that the interpretation of the chemical potential as the Gibbs free energy per particle holds only for the case of simple fluids—in the case of a mixture of several chemical species, it is no longer valid.

If we take the differential of E via the Euler equation, we obtain

$$\mathrm{d}E = T\mathrm{d}S + S\mathrm{d}T + \sum_{i=1}^{r} (f_i \mathrm{d}x + X_i \mathrm{d}f_i). \tag{2.126}$$

By subtracting both sides of this equation from the expression of $\mathrm{d}E$ given in eq. (2.61), we obtain the **Gibbs-Duhem equation**:

$$S\mathrm{d}T + \sum_{i=1}^{r} X_i \mathrm{d}f_i = 0. \tag{2.127}$$

Dividing both sides of this equation by any (nonvanishing) extensive variable, we realize that it establishes a relation between intensive variables and densities. It therefore plays a key role in the interpretation of thermodynamic data. For example, in the case of simple fluids one arrives at

$$S\mathrm{d}T - V\mathrm{d}p + N\mathrm{d}\mu = 0. \tag{2.128}$$

By dividing with respect to the number of particles N, one obtains the Gibbs-Duhem equation in the form

$$\mathrm{d}\mu = v\,\mathrm{d}p - s\,\mathrm{d}T, \tag{2.129}$$

where v represents the volume per particle and s the entropy per particle. Therefore, if we know the equation of state $v = v(p, T)$, we can integrate the equation for $T = \text{const.}$, starting, say, from $p = p_0$:

$$\mu(p, T) = \mu(p_0, T) + \int_{p_0}^{p} \mathrm{d}p'\, v(p', T). \tag{2.130}$$

To determine $\mu(p_0, T)$ it is necessary to know the value of the entropy per particle s for $p = p_0$ and all values of T:

$$\mu(p_0, T) = \mu(p_0, T_0) - \int_{T_0}^{T} \mathrm{d}T'\, s(p_0, T'). \tag{2.131}$$

In this way the chemical potential μ can be evaluated up to an additive constant.

2.17 Magnetic Systems

In statistical mechanics, one is often dealing with systems whose thermodynamic variables include the total magnetic moment (usually denoted by $\boldsymbol{M}$). It is an extensive quantity that can therefore be added to the other extensive variables S, V, N, and so on.

Let us suppose that we have a sample in the form of a narrow cylinder, placed inside a solenoid, identified by a number of coils per unit of length equal to ν, inside which a current I is circulating. In this case, the sample is subjected to a uniform magnetic field $\boldsymbol{H}$—oriented so that it is parallel to the cylinder's axis—and whose amplitude is νI. Within the geometry we are examining, if the cylinder is homogeneous, it develops a uniform magnetization, with a certain value $\boldsymbol{m}$ of magnetic moment density per unit of volume. As a consequence, the magnetic induction vector $\boldsymbol{B}$ is uniform, parallel to $\boldsymbol{H}$, and takes the form $\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{m})$, where $\mu_0 \simeq 4\pi \cdot 10^{-7}\,\mathrm{H\,m^{-1}}$ is the magnetic permeability of the

vacuum. (The actual values of this and of other fundamental physical constants are listed in sec. H.)

Let us now suppose that we vary the current circulating in the solenoid by ΔI. The work performed on the system and associated with this variation is given by

$$đW = -\int_{t_{\text{in}}}^{t_{\text{fin}}} \mathrm{d}t\, I(t) \nu_{\text{ind}}, \tag{2.132}$$

where ν_{ind} is the induced electromotive force. This force is expressed in terms of the flux Φ of the magnetic induction vector $\boldsymbol{B}$ by

$$\nu_{\text{ind}} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}. \tag{2.133}$$

We thus obtain

$$đW = -\int_{t_{\text{in}}}^{t_{\text{fin}}} \mathrm{d}t\, I(t) \frac{\mathrm{d}\Phi(t)}{\mathrm{d}t} \simeq I\Delta\Phi, \tag{2.134}$$

where we have neglected higher-order contributions in ΔI. The value of the concatenated flux Φ is equal to $B\nu\ell A = \nu BV$ for a solenoid of length ℓ and cross section A, and therefore of volume $V = A\ell$. We thus obtain

$$\Delta\Phi = \nu V \Delta B = \nu V \mu_0 (\Delta H + \Delta M) = \nu V \mu_0 H \Delta H + \nu \mu_0 H \Delta M. \tag{2.135}$$

As a consequence, the work $đW$ can be expressed by

$$đW = \nu V \mu_0 I \Delta H + \nu \mu_0 I \Delta M = \mu_0 V H \Delta H + \mu_0 H \Delta M. \tag{2.136}$$

The first term represents the variation in the energy of the magnetic field due to variations in the magnetic field itself. In fact, this energy is equal to

$$E_{\text{magn}} = \int \mathrm{d}\boldsymbol{r}\, \frac{1}{2}\mu_0 H^2, \tag{2.137}$$

and the value of its variation is therefore $\Delta E_{\text{magn}} = \mu_0 \int \mathrm{d}\boldsymbol{r}\, \boldsymbol{H} \cdot \Delta \boldsymbol{H}$. The second term represents the work performed on the sample in order to change its magnetization value. The contribution of the extensive variable $\boldsymbol{M}$ to "mechanical" work is therefore given by

$$đW_{\text{magn}} = \mu_0 \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{M}. \tag{2.138}$$

We can see that the variable conjugate to magnetization (in other words, to the total magnetic dipole) is the *magnetic field* (up to the factor μ_0).

Since dipolar interactions (which occur between regions that are magnetically polarized) are long range, a certain caution is necessary when dealing with the magnetization $\boldsymbol{M}$ as an ordinary thermodynamic variable. More specifically, given equal amounts of material and volume, samples with different shapes possess different energies even if immersed in the same magnetic field. One usually accounts for this effect by introducing a **demagnetization factor**, which depends on shape.

It is also useful to introduce a thermodynamic potential related to the Legendre transform of E with respect to $\boldsymbol{M}$:

$$\mathcal{E}(S, \boldsymbol{H}, X_1, \ldots, X_r) = E(S, \boldsymbol{H}, X_1, \ldots, X_r) - \mu_0 \boldsymbol{H} \cdot \boldsymbol{M}. \tag{2.139}$$

The differential of this thermodynamic potential has the expression

$$\mathrm{d}\mathcal{E} = T\mathrm{d}S - \mu_0 \boldsymbol{M} \cdot \mathrm{d}\boldsymbol{H} + \sum_{i=1}^{r} f_i \mathrm{d}X_i. \tag{2.140}$$

One should note that, with our conventions, there are no extra signs to be applied to the differentials.

2.18 Equations of State

Relations between the densities and the intensive variables obtained by taking derivatives of the fundamental equation (in one of its various representations) are called **equations of state**.

If, for example, we consider the Gibbs free energy for a simple fluid, we arrive at

$$V = \frac{\partial G}{\partial p}\bigg)_{T,N}. \tag{2.141}$$

In the case of the simple fluid, we have another independent equation of state:

$$S = -\frac{\partial G}{\partial T}\bigg)_{p,N}, \tag{2.142}$$

which expresses the entropy as a function of p, T, and N. Actually, the two equations of state are not completely independent, because of the Maxwell relation

$$\frac{\partial V}{\partial T}\bigg)_{p} = -\frac{\partial S}{\partial p}\bigg)_{p}. \tag{2.143}$$

Indeed, by integrating the first equation of state, we obtain the entropy as a function of p and T, up to an additive term independent of p:

$$\begin{aligned} S(p, T, N) &= \int_{(p_0,T_0)}^{(p,T)} \left[\frac{\partial S}{\partial p}\bigg)_{T,N} \mathrm{d}p + \frac{\partial S}{\partial T}\bigg)_{p,N} \mathrm{d}T \right] + S(p_0, T_0, N) \\ &= \left[-\int_{(p_0,T)}^{(p,T)} \frac{\partial V}{\partial T}\bigg)_{p} \mathrm{d}p + \int_{(p_0,T_0)}^{(p_0,T)} \frac{C_p \mathrm{d}T}{T} \right] + S(p_0, T_0, N). \end{aligned} \tag{2.144}$$

We have used the Maxwell relation (2.143) and the relation

$$\frac{\partial S}{\partial T}\bigg)_{p,N} = \frac{C_p}{T}, \tag{2.145}$$

where C_p is the heat capacity at constant pressure. The first term of this expression can therefore be obtained from the equation of state; however, an independent measurement is necessary to unambiguously evaluate the second term. And vice versa, Maxwell relations allow one to check the consistency of thermodynamic data. Due to the extensivity of G, both V and S are proportional to N. The corresponding densities, $v = V/N$ and $s = S/N$, are therefore independent of the system size.

The **ideal gas** is a simple fluid that satisfies the equation of state

$$p = \frac{Nk_B T}{V}, \tag{2.146}$$

where $k_B \simeq 1.381 \cdot 10^{-23}\,\mathrm{JK}^{-1}$ is **Boltzmann's constant**.

Maxwell's relations allow us to prove that in an ideal gas, the internal energy depends only on N and T (and not on the volume V). One has in fact

$$\left.\frac{\partial E}{\partial V}\right)_T = \left.\frac{\partial E}{\partial V}\right)_S + \left.\frac{\partial E}{\partial S}\right)_V \left.\frac{\partial S}{\partial V}\right)_T = -p + T \left.\frac{\partial p}{\partial T}\right)_V = 0. \tag{2.147}$$

Moreover, the entropy of an ideal gas is the sum of a term that depends only on T and one that depends only on V. In fact, one gets

$$\left.\frac{\partial S}{\partial V}\right)_T = \left.\frac{\partial p}{\partial T}\right)_V = \frac{Nk_B}{T}, \tag{2.148}$$

and therefore

$$\frac{\partial^2 S}{\partial T \partial V} = 0, \tag{2.149}$$

which is what we intended to prove.

2.19 Stability

Let us now consider two systems, 1 and 2, in contact with each other, which can exchange energy, particles, and volume. Following the line of reasoning pursued in section 2.5, we must have at equilibrium

$$\begin{aligned} T^{(1)} &= T^{(2)}, \\ p^{(1)} &= p^{(2)}, \\ \mu^{(1)} &= \mu^{(2)}. \end{aligned} \tag{2.150}$$

Continuing to follow the same line of reasoning from section 2.5, we can conclude that if the other intensive qualities are the same in the two systems, then

1. if initially $T^{(1)} > T^{(2)}$, system 1 releases energy to system 2;
2. if initially $p^{(1)} > p^{(2)}$, system 1 expands at the expense of system 2—in other words, system 2 releases volume to system 1;
3. if initially $\mu^{(1)} > \mu^{(2)}$, system 1 releases particles to system 2.

These results are due to the convexity of the surface $E(S, V, \mu)$.

The convexity is connected to the stability of thermodynamic equilibrium. Stability actually implies that the thermodynamic potential cannot decrease in the course of any virtual displacement. Let us consider an infinitesimal virtual variation δX_i of the extensive variable X_i, and let us suppose that the intensive variables $f_0, \ldots, f_{i-1}$ and the extensive variables $X_{i+1}, \ldots, X_r$ are kept fixed. Let us denote the corresponding thermodynamic potential as follows:

$$\Phi(f_0 \ldots, f_{i-1}, X_i, X_{i+1}, \ldots, X_r) = E - \sum_{i=0}^{i-1} f_i X_i. \tag{2.151}$$

One then obtains

$$\delta^2 \Phi = \frac{1}{2} \left. \frac{\partial^2 \Phi}{\partial X_i^2} \right)_{f_0, \ldots, f_{i-1}, X_{i+1}, \ldots, X_r} \delta X_i^2 \geq 0. \tag{2.152}$$

In particular, the specific heat is always nonnegative, whatever constraints are imposed on the system:

$$C = T \frac{\partial S}{\partial T} \geq 0. \tag{2.153}$$

If this were not the case, a small fluctuation in temperature that, for example, makes colder one region of the system would lead to it claiming heat from the surrounding regions. With negative specific heat, this energy would make the temperature of the region in question diminish further, and in this manner, the energy fluctuation would be amplified.

Moreover, the specific heat with a fixed intensive variable f_i is always greater than that with the corresponding fixed extensive variable. In fact, one has

$$C_{f_i} = T \left. \frac{\partial S}{\partial T} \right)_{f_i} = T \left. \frac{\partial S}{\partial T} \right)_{X_i} + T \left. \frac{\partial S}{\partial X_i} \right)_T \left. \frac{\partial X_i}{\partial T} \right)_{f_i}. \tag{2.154}$$

We now apply the Maxwell relation

$$\left. \frac{\partial S}{\partial X_i} \right)_T = - \left. \frac{\partial f_i}{\partial T} \right)_{X_i}, \tag{2.155}$$

and the following identity (cf. sec. C), which holds when x, y, and z satisfy a relation of the kind $f(x, y, z) = 0$:

$$\left. \frac{\partial x}{\partial y} \right)_z = - \left. \frac{\partial x}{\partial z} \right)_y \left. \frac{\partial z}{\partial y} \right)_x. \tag{2.156}$$

We thus obtain

$$C_{f_i} = T \left. \frac{\partial S}{\partial T} \right)_{X_i} + T \left. \frac{\partial f_i}{\partial T} \right)_{x_i}^2 \left. \frac{\partial X_i}{\partial f_i} \right)_T \geq T \left. \frac{\partial S}{\partial T} \right)_{X_i}. \tag{2.157}$$

In particular, the specific heat at constant pressure C_p is always larger than the specific heat at constant volume C_V. For an ideal gas one has **Meyer's relation**:

$$C_p - C_V = Nk_{\mathrm{B}}. \tag{2.158}$$

One can show that in an infinitesimal transformation at constant entropy one has

$$\frac{\partial p}{\partial V}\Big)_S = -\gamma_C \frac{p}{V}, \tag{2.159}$$

where $\gamma_C = C_p/C_V$. Assuming $\gamma_C = \text{const.}$, one can integrate this relation, and obtain the **law of adiabatics**:

$$pV^{\gamma_C} = \text{const.} \tag{2.160}$$

The derivation of these results is left as an exercise.

Exercise 2.9 (Law of adiabatic processes) Prove eqs. (2.158), (2.159), and (2.160). Deduce that in a reversible adiabatic process one has

$$pT^{\gamma_C/(1-\gamma_C)} = \text{const.},$$

provided $\gamma_C = \text{const.}$

Exercise 2.10 (Reech's relation) Show that the ratio $\gamma_C = C_p/C_V$ is equal to the ratio κ_T/κ_S of the isothermal compressibility

$$\kappa_T = -\frac{1}{V}\frac{\partial V}{\partial p}\Big)_T$$

to the adiabatic compressibility

$$\kappa_T = -\frac{1}{V}\frac{\partial V}{\partial p}\Big)_S.$$

Exercise 2.11 (Adiabatic atmosphere) It is a well-known fact that atmospheric temperature decreases with an increase in elevation. As a first approximation, given the low thermal conductivity of air, we can assume that the atmosphere is at equilibrium with respect to vertical adiabatic shifts of small quantities of air. By using the equation of state for ideal gases, the law of adiabatic processes, and Stevin's law.

$$\frac{\partial p}{\partial z} = -m_{\mathrm{mol}}\, g \frac{N}{V},$$

where m_{mol} is the average molecular mass of air and g the gravitational acceleration, prove that the thermal gradient of the atmosphere is given by

$$\frac{\partial T}{\partial z}\Big)_S = -\frac{m_{\mathrm{mol}}\, g}{k_{\mathrm{B}}}\left(1 - \gamma_C^{-1}\right). \tag{2.161}$$

Evaluate this quantity numerically. We recall that for air, $\gamma_C = C_p/C_V = 7/5$.

Exercise 2.12 Let us consider a container whose walls are adiabatic, in which there is an opening closed by a faucet. A void is created inside the container, which is then placed in contact with the atmosphere at pressure p_0 and temperature T_0. At this point, the faucet is opened and air rapidly enters the container. After a brief interval, internal pressure becomes equal to the external pressure. Calculate the temperature of the air inside the container under these conditions.

2.20 Chemical Reactions

Let us now consider a mixture of r chemical species, $A_1, \ldots, A_r$, which can be transformed into one other by a reaction of the following type:

$$\nu_1 A_1 + \cdots + \nu_k A_k \leftrightharpoons \nu_{k+1} A_{k+1} + \cdots + \nu_r A_r. \tag{2.162}$$

As is well known, this formula states that one takes ν_1 molecules of the species A_1, ν_2 of species $A_2, \ldots, \nu_k$ of species A_k (the **reactants**), and one produces ν_{k+1} molecules of species $A_{k+1}, \ldots, \nu_r$ of species A_r (the **products**). The integers $\nu_1, \ldots, \nu_r$ are called **stoichiometric coefficients**.

There is no fundamental distinction between reactants and products, because each reaction can go both in one direction and in the reverse, as is indicated by the two arrows in the formula. We can conventionally assign negative stoichiometric coefficients to the products, so as to write this formula as a formal equation:

$$\sum_{i=1}^{r} \nu_i A_i = 0. \tag{2.163}$$

The idea of this representation is that, if the reaction takes place, at each moment of the variation, δN_i of the number of particles of species i is proportional to the corresponding stoichiometric coefficient, taken with its own sign:

$$\frac{\delta N_1}{\nu_1} = \frac{\Delta N_2}{\nu_2} = \cdots = \frac{\Delta N_r}{\nu_r}. \tag{2.164}$$

Obviously, the variation will be negative for the reactants and positive for the products if the reaction goes from left to right, and vice versa.

If temperature and pressure are kept fixed, we can calculate the variation of Gibbs free energy for a certain variation in the number of particles due to the reaction:

$$\delta G = \sum_i \frac{\partial G}{\partial N_i}\bigg)_{T,p} \delta N_i \propto \sum_i \frac{\partial G}{\partial N_i}\bigg)_{T,p} \nu_i = \sum_i \mu_i \nu_i. \tag{2.165}$$

Since at equilibrium one must have $\delta G = 0$ for any virtual variation of the N_i, one has

$$\sum_i \nu_i \mu_i = 0, \tag{2.166}$$

where the stoichiometric coefficients ν_i are taken with their corresponding signs. We have thus obtained the equilibrium criterion with respect to the reaction described by eq. (2.162).

2.21 Phase Coexistence

It frequently happens that two systems characterized by different thermodynamic density values can maintain thermodynamic equilibrium even in the absence of constraints on the mutual exchange of extensive quantities. This situation is called **phase coexistence**. In the case of a simple fluid, it is realized, for example, when a liquid coexists with its vapor inside a container. In this case, the intensive variables assume the same value in both systems, while densities (for example, the volume v per particle, given by $v = V/N$) assume different values. We refer to each such coexisting homogeneous system as a **phase**.

One can describe phase coexistence by saying that the equation of state

$$v = v(T, p) \tag{2.167}$$

does not admit of a unique solution, but instead allows for at least the two solutions, $v = v_{\text{liq}}$ and $v = v_{\text{vap}}$, which correspond to the liquid and the vapor, respectively.

The postulates of thermodynamics then imply that *all* intermediate values of v between v_{liq} and v_{vap} satisfy the equation of state. In effect, since the liquid and vapor coexist and can exchange particles, the chemical potential of the liquid has to be equal to that of the vapor:

$$\mu_{\text{liq}}(T, p) = \mu_{\text{vap}}(T, p). \tag{2.168}$$

On the other hand, we know that for a simple fluid, the chemical potential is equal to the Gibbs free energy per particle:

$$\mu = \frac{G}{N}. \tag{2.169}$$

The Gibbs free energy is the thermodynamic potential that describes equilibrium when pressure p and temperature T are fixed, as a function of the number of particles N. Therefore, in the situation we are considering, the Gibbs free energy in the total system does not depend on the number of particles N_{liq} and N_{vap} that make up the liquid system and the vapor system, respectively:

$$G = G_{\text{liq}} + G_{\text{vap}} = N_{\text{liq}}\mu_{\text{liq}} + N_{\text{vap}}\mu_{\text{vap}} = \left(N_{\text{liq}} + N_{\text{vap}}\right)\mu = N\mu. \tag{2.170}$$

The volume per particle of a system in which N_{liq} particles are part of the liquid and N_{vap} particles are part of the vapor is given by

$$v = \frac{V}{N} = \frac{V_{\text{liq}} + V_{\text{vap}}}{N} = \frac{N_{\text{liq}}v_{\text{liq}} + N_{\text{vap}}v_{\text{vap}}}{N}. \tag{2.171}$$

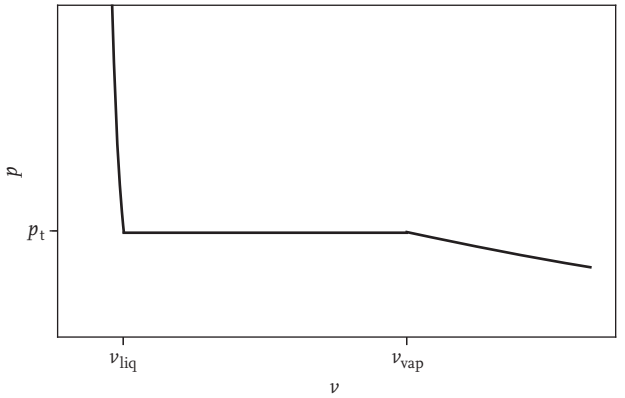


FIGURE 2.8. Isotherm p as a function of the volume per particle v for a simple fluid. The phase coexistence is indicated by the horizontal segment for $p = p_t$ included between v_{liq} and v_{vap}.

We thus obtain

$$v = x_{liq} v_{liq} + x_{vap} v_{vap}, \tag{2.172}$$

where $x_{liq} = N_{liq}/N$ is the fraction of particles in the liquid and $x_{vap} = 1 - x_{liq}$ that of the particles in the vapor. As a consequence, the value of v lies somewhere between v_{liq} and v_{vap}. Obviously, an analogous result obtains for the density $\rho = 1/v$ and for the other densities one could define.

We now discuss different aspects of phase coexistence in a simple fluid.

In the equation of state $p = p(T, v)$, the discussion we just completed shows that phase coexistence appears as a horizontal segment, for a given value of T, and for values of v included between v_{liq} and v_{vap} (see figure 2.8). Let us now consider the Helmholtz free energy F. The pressure p is obtained as the derivative of F with respect to V:

$$p = -\left.\frac{\partial F}{\partial V}\right)_{T,X}. \tag{2.173}$$

Phase coexistence therefore manifests itself because, for a given value of T, and other extensive variables X keeping a fixed value, the curve $f(v) = F(T, V)/N$ exhibits a straight segment with slope $-p_t$, for v between v_{liq} and v_{vap} (see figure 2.9, left panel).

The properties of the Legendre transform (see appendix A) imply that the Gibbs free energy, and therefore the chemical potential μ, has an angular point at $p = p_t$, as a function of p at a fixed value of T (see figure 2.9, right panel). The two slopes that coexist at the angular point correspond to v_{liq} and v_{vap}, respectively. Since $\mu = G/N$, the shape of the isotherm $\mu = \mu(p)$ is the same as that of G. The concavity of G vs. p implies that the smaller specific volume (and therefore the smaller derivative) always corresponds to the phase that is stable for larger values of p.

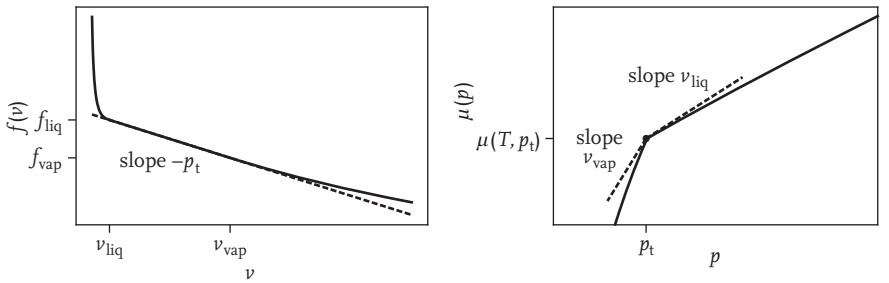


FIGURE 2.9. Left: Isotherm of the Helmholtz free energy per particle, $f = F/N$, as a function of volume per particle $v = V/N$. The straight line of slope $-p_t$ corresponds to phase coexistence. Right: Isotherm of the Gibbs free energy per particle: $\mu = G/N$ as a function of pressure p. The angular point at p_t corresponds to phase coexistence.

2.22 The Clausius-Clapeyron Equation

If we consider a thermodynamic system as a function of its intensive variables, we can identify some regions in which the thermodynamic properties vary regularly with variations of their values. These regions represent thermodynamically stable phases and are bounded by curves that represent **phase transitions**. The phase transitions can be **discontinuous**, like the phase coexistence we just discussed, or **continuous**. In the first case, the densities present a discontinuity at the transition, while in the second, they vary with continuity, even though their derivatives can exhibit some singularities. One often also employs the following terminology: discontinuous transitions are called **first-order transitions**, while continuous ones are called **second-order transitions**.

In the case of a simple fluid, it is possible to identify the transition curve within the plane of the intensive variables (p, T)—in other words, the curve $p = p_t(T)$—from the condition of equality of the chemical potential μ between the two coexisting phases:

$$\mu_{\text{liq}}(T, p_t(T)) = \mu_{\text{vap}}(T, p_t(T)). \tag{2.174}$$

The pressure $p_t(T)$ at which the transition occurs is also called **vapor pressure**. It is possible to relate this curve locally with the discontinuity of densities at transition. To obtain this relation, let us take the total derivative of this equation with respect to T, along the transition line $p = p_t(T)$. We obtain

$$\left(\frac{\partial \mu_{\text{liq}}}{\partial p}\right)_T \left(\frac{\mathrm{d}p_t}{\mathrm{d}T} + \frac{\partial \mu_{\text{liq}}}{\partial T}\right)_p = \left(\frac{\partial \mu_{\text{vap}}}{\partial p}\right)_T \left(\frac{\mathrm{d}p_t}{\mathrm{d}T} + \frac{\partial \mu_{\text{vap}}}{\partial T}\right)_p. \tag{2.175}$$

Therefore, by the Gibbs-Duhem equation, we have

$$v_{\text{liq}} \frac{\mathrm{d}p_t}{\mathrm{d}T} - s_{\text{liq}} = v_{\text{vap}} \frac{\mathrm{d}p_{\text{vap}}}{\mathrm{d}T} - s_{\text{vap}}, \tag{2.176}$$

where v and s are respectively the volume and the entropy per molecule of the two coexisting phases. We thus obtain

$$\frac{\mathrm{d}p_t}{\mathrm{d}T} = \frac{s_{\text{vap}} - s_{\text{liq}}}{v_{\text{vap}} - v_{\text{liq}}}. \tag{2.177}$$

This equation, which can be applied to each case of phase coexistence, is called the **Clausius-Clapeyron equation**. We can see that under usual conditions, since $v_{vap} > v_{liq}$ and $s_{vap} = s_{liq} + Q_{evap}/T > s_{liq}$ (where $Q_{evap} > 0$ is the latent heat of evaporation), the vapor pressure $p_t(T)$ increases with the temperature:

$$\frac{dp_t}{dT} > 0. \tag{2.178}$$

In the case of the melting of water, in which, as is well known, the phase at low temperature (ice) has a volume per particle that is larger than the phase at higher temperature (liquid), we have

$$s_{liq} > s_{ice}; \tag{2.179}$$

$$v_{liq} < v_{ice}. \tag{2.180}$$

The Clausius-Clapeyron equation therefore predicts that

$$\frac{dp_t}{dT} < 0. \tag{2.181}$$

The curve of the melting pressure versus temperature for water at $T = 0°C$ has therefore a negative slope. In other words, when exerting a pressure on ice at 0°C, it melts. This effect is fairly reasonable, since it is a manifestation of stability—given an increase in pressure, the system reacts by assuming the phase that occupies less volume.

2.23 The Coexistence Curve

We can represent the phase diagram in the plane (v, T), in which the intensive variable T is accompanied by the volume per particle v, rather than in the plane (p, T) of the intensive variables. In this manner, phase coexistence is represented by the existence of a forbidden region $v_{liq}(T) < v < v_{vap}(T)$ in the plane. Outside this region, it is possible to obtain any given value of v in a homogeneous system. Within this region, instead, the system separates into a liquid phase and a vapor phase (see figure 2.10). The fraction x_{liq} of particles in the liquid phase (and the analogous fraction in the vapor phase) is determined by the condition that volume per particle in the whole system is equal to v. One thus obtains

$$x_{liq} = \frac{v_{vap} - v}{v_{vap} - v_{liq}}; \qquad x_{vap} = \frac{v - v_{liq}}{v_{vap} - v_{liq}} = 1 - x_{liq}. \tag{2.182}$$

This result is known as the **lever rule**. In effect, if one considers the segment included between v_{liq} and v_{vap} as the arm of a lever, with its fulcrum corresponding to v, the fractions x_{liq} and x_{vap} are proportional to the weights one would need to place at each extremity so that the lever would be in equilibrium.

Let us note that while the coexistence curve is defined without ambiguities in the plane (v, T), because the points that represent the phases in coexistence must correspond to the same value of T (an intensive variable), this situation is no longer valid if we place ourselves in a plane characterized by two densities—for instance, (v, s) (where s is the entropy per

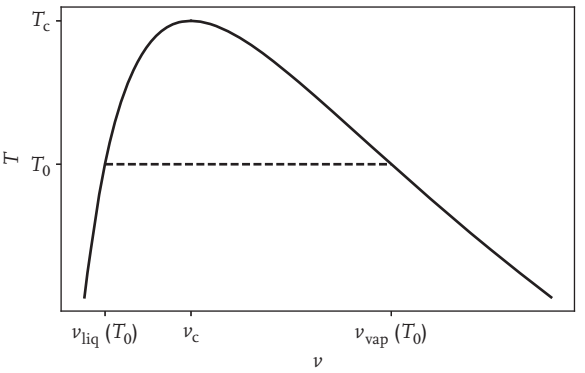


FIGURE 2.10. Coexistence curve for a simple fluid. The temperature T is on the y axis, and the volume per particle v is on the x axis. The critical point corresponds to (v_c, T_c). The horizontal line represents the possible values of v for a system at phase coexistence at temperature T_0.

particle). In this case, the coexistence of two phases occurs along straight segments that connect the points (v_{liq}, s_{liq}) and (v_{vap}, s_{vap}), which correspond to equal temperature values. These segments must be represented explicitly, and are called **tie lines**.

2.24 Coexistence of Several Phases

It is also possible for several phases to coexist for certain values of the intensive variables. For example, in the case of a simple fluid, one can observe the coexistence of a solid, a liquid, and a vapor for specific values of the p_t and of the temperature value T_t. This point is called the **triple point**. At the triple point, the chemical potentials corresponding to the solid, liquid, and vapor phases must be equal:

$$\mu_{sol}(T_t, p_t) = \mu_{liq}(T_t, p_t) = \mu_{vap}(T_t, p_t). \tag{2.183}$$

Here we have *two* equations in the two intensive variables p and T, instead of the single equation that we had obtained for the coexistence of the liquid and the vapor. For a simple fluid, therefore, the coexistence of three phases occurs only for isolated points of the plane (p, T). There can be several triple points, corresponding to the different possible solid phases. The most familiar example is the triple point of water—$T_t = 273.16$ K; $p_t = 0.6116$ kPa—which was used until 2019 to define the temperature unit (1 degree Kelvin).

Let us now consider a more general case—in other words, a mixture of particles belonging to r different chemical species. Let us suppose that we are looking for the coexistence of q phases ($q \geq 2$). At equilibrium, all the intensive variables must assume the same value in the coexisting phases. We will therefore have a specific value for the pressure and the temperature, and in addition the chemical potential of each species will have to assume the same value in all the different phases. If we denote the chemical potential of

species i in phase α as $\mu_i^{(\alpha)}$, we will then have

$$\mu_i^{(\alpha)} = \mu_i, \qquad i = 1, \ldots, r, \quad \alpha = 1, \ldots, q. \tag{2.184}$$

In this equation, μ_i is the shared value taken by the chemical potential of species i. We thus obtain $r(q-1)$ equations for $q(r-1)+2$ unknown values. These unknown values are p, T, and the $q(r-1)$ independent densities $x_i^{(\alpha)}$ of species i in phase α. Generically speaking, therefore, $f = 2 - q + r$ free parameters remain. For $f = 0$, coexistence will occur in isolated points of the phase diagram; for $f = 1$, along a line; and so on. The quantity f is called **variance**. The reasoning we have followed therefore allows us to conclude that

$$f = 2 - q + r, \tag{2.185}$$

which is called the **Gibbs phase rule**.

The Gibbs phase rule must be used with caution, because cases in which it appears to be violated are not uncommon. A classic example is the coexistence of two solid phases of carbon (graphite and diamond) at ambient temperature. Logically, they should coexist along a coexistence line, and therefore, once the temperature has been fixed, for a specific value of pressure. They instead appear to coexist at any pressure (as long as it is low) and even in a "vacuum"—in other words, when their vapor is present. The solution to this paradox lies in the fact that the transformation of one phase into the other at ambient temperature is extremely slow. Thermodynamically speaking, the phase that is stable at ambient temperature and pressure is graphite; however, as is well known, the time necessary for a diamond (even a very small one) to spontaneously turn into graphite is of the order of millions of years.

2.25 The Critical Point

Let us once again consider the simple fluid and observe that the coexistence of liquid and vapor cannot be obtained for temperatures higher than a certain temperature T_c, called the **critical temperature**. To be more precise, the transition curve ends at a point (p_c, T_c), where p_c is the critical pressure. For $T < T_c$, the difference $v_{vap} - v_{liq}$ approaches zero continuously when T gets close to T_c—the discontinuity of thermodynamic densities tends to vanish, or in other words, the transition goes from being discontinuous to being continuous (and finally to disappearing at higher temperatures).

The critical point is a thermodynamic state with exceptional properties. For example, since for $T < T_c$ and $p = p_t(T)$ one gets $\partial p / \partial V)_T = 0$ (within the coexistence curve); this relation must ultimately be valid also at the critical point—in other words, the system's compressibility diverges:

$$\chi = -\frac{1}{V}\left.\frac{\partial V}{\partial p}\right)_T \to \infty, \qquad \text{for } T \to T_c^+. \tag{2.186}$$

Various thermodynamic properties exhibit analogous singularities.

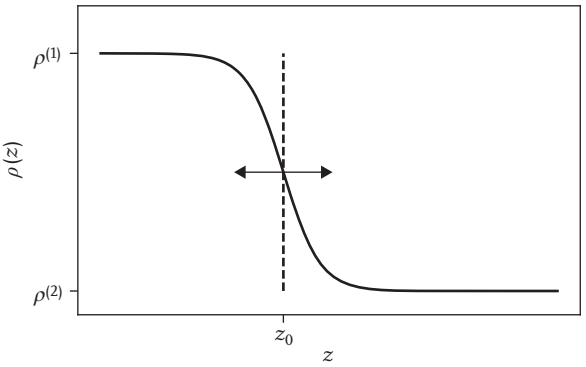


FIGURE 2.11. Local density profile in a liquid-vapor interface: ρ as a function of height z. The dashed line is placed at the location z_0 obtained by Gibbs's condition. The arrow marks the interface width.

The study of the thermodynamic properties of the critical point cannot proceed very far without appealing to genuine statistical mechanics. For this reason, we will discuss these properties at the appropriate moment. In addition to the critical point in the liquid-gas transition, there are many examples of analogous situations (the Curie point in ferromagnets, the λ point in liquid helium, and so on). They form the simplest class of continuous transitions.

2.26 Planar Interfaces

An **interface** forms between two different phases in equilibrium. The contributions of the interface to thermodynamic properties are negligible for large systems (because they are proportional to the system's surface, not its volume), but they are interesting by themselves.

Let us consider a container in which the liquid and vapor phases of a simple fluid coexist. If the container is large enough, the interface is horizontal for the greater part of the container's cross section. We can measure the local density of the fluid as a function of height z and thus obtain a curve like the one shown in figure 2.11. At great distances from the interface, the density is constant and equal to the density of the liquid or the vapor. However, if the distance from the interface is small (but still very large when compared to atomic scales), the density varies with continuity. We can therefore introduce the notion of **interface width**.

From the thermodynamic point of view, however, it is useful to be able to consider the interface as a geometric surface, placed at a certain height z_0. Let us consider the system composed of phases 1 and 2, with the addition of an interface, as shown in figure 2.12. The internal energy is given by the sum of the contributions from each phase, to which is then added the energy of the interface:

$$E = E^{(1)} + E^{(2)} + E^{\text{int}}. \tag{2.187}$$

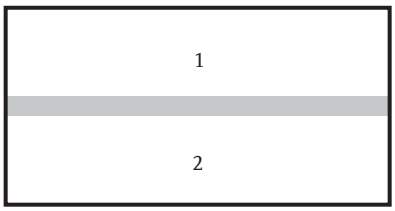


FIGURE 2.12. Coexistence of two phases, with their interface.

The contribution of the interface depends on its area A. Let us then write the differential of E in the form

$$\mathrm{d}E = T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N + \gamma\,\mathrm{d}A. \tag{2.188}$$

The quantity

$$\gamma = \frac{\partial E}{\partial A}\Big)_{S,V,N} \tag{2.189}$$

is called **surface tension**. It must be positive—otherwise, the interface would tend to increase its area, filling all space available, and we could no longer describe the system as the coexistence of two distinct phases. (Situations of this kind occur in microemulsions.)

As was the case with internal energy, the other extensive variables can also be obtained by adding the contributions of the two phases to that of the interface. By using Euler's equation, we arrive at

$$\begin{aligned} E = {} & T\left(S^{(1)} + S^{(2)} + S^{\mathrm{int}}\right) \\ & - p\left(V^{(1)} + V^{(2)} + V^{\mathrm{int}}\right) + \mu\left(N^{(1)} + N^{(2)} + N^{\mathrm{int}}\right) + \gamma A. \end{aligned} \tag{2.190}$$

If we want to treat the interface as a geometric surface, we must set $V^{\mathrm{int}} = 0$. Moreover, if we choose the height z_0 of the interface in a suitable fashion, we can arrange things so that

$$\int_{-\infty}^{z_0} \mathrm{d}z\,\left[\rho^{(1)} - \rho(z)\right] = \int_{z_0}^{+\infty} \mathrm{d}z\,\left[\rho(z) - \rho^{(2)}\right]. \tag{2.191}$$

This condition is equivalent to imposing $N^{\mathrm{int}} = 0$ and is known as the **Gibbs condition**. We thus obtain

$$E^{\mathrm{int}} = TS^{\mathrm{int}} + \gamma\, A. \tag{2.192}$$

Due to extensivity, S^{int} must also be proportional to the area A of the interface. From this equation, we can obtain the physical interpretation of γ:

$$\gamma = \frac{E^{\mathrm{int}} - TS^{\mathrm{int}}}{A}. \tag{2.193}$$

Therefore, γ is equal to the Helmholtz free energy per unit interface area.

Let us now assume that we are dealing with a system with two coexisting phases, but where several chemical species are present. We can still define the position of the interface

so as to have $N^{\text{int}} = 0$ for species 1 (which we will assume is prevalent—the solvent). In general, however, we will have to assign a certain quantity of particles of the other species to the interface. We will then have

$$\mathrm{d}E^{\text{int}} = T\mathrm{d}S^{\text{int}} + \sum_{i=2}^{r} \mu_i \mathrm{d}N_i^{\text{int}} + \gamma\, A. \tag{2.194}$$

In the case of only two species, we will have

$$\mathrm{d}E^{\text{int}} = T\mathrm{d}S^{\text{int}} + \mu_2 \mathrm{d}N_2^{\text{int}} + \gamma\, \mathrm{d}A, \tag{2.195}$$

and as a result of Euler's equation,

$$E^{\text{int}} = TS^{\text{int}} + \mu_2 N_2^{\text{int}} + \gamma\, A. \tag{2.196}$$

We thus obtain the analog of the Gibbs-Duhem equation for the surface tension:

$$\mathrm{d}\gamma = -s^{\text{int}}\mathrm{d}T - \rho_2^{\text{int}}\mathrm{d}\mu_2, \tag{2.197}$$

where s^{int} is the entropy per unit area and ρ_2^{int} the surface density per unit area of species 2.

At constant temperature T, we have the equation

$$\left.\frac{\partial \gamma}{\partial \mu_2}\right)_T = -\rho_2^{\text{int}}, \tag{2.198}$$

whose solution is called the **Gibbs isotherm**. Let us note that one has

$$0 < \rho_2^{\text{int}} = -\left.\frac{\partial \gamma}{\partial \mu_2}\right)_T = -\left.\frac{\partial \gamma}{\partial \rho_2^{\text{int}}}\right)_T \left.\frac{\partial \rho_2^{\text{int}}}{\partial \mu_2}\right)_T. \tag{2.199}$$

The last factor is positive due to stability. We thus obtain

$$\left.\frac{\partial \gamma}{\partial \rho_2^{\text{int}}}\right)_T < 0, \tag{2.200}$$

which implies that surface tension decreases when a solute accumulates on the interface. This effect is used to decrease the surface tension and favor mixing between liquids, by introducing particles that accumulate at the interface (the **surfactants** or **tensioactives**).

Exercise 2.13 (Theory of elasticity) Consider an elastic band with a fixed cross section as a thermodynamic system. The thermodynamic variables that characterize it within the energy scheme are its entropy S, length L, and a measure N of its mass, all assumed to be extensive. The expression of $\mathrm{d}E$ is given by

$$\mathrm{d}E = T\,\mathrm{d}S + f\,\mathrm{d}L + \mu\,\mathrm{d}N,$$

where f is the tension acting on the band, and μ is the chemical potential.

One observes experimentally that the tension f, when the length L and the mass m are constant, increases proportionally to temperature. Thus the equation of state is given by

$$f = \kappa(L, N)\, T,$$

where κ depends on L and N. Hooke's law, on the other hand, implies that the relative lengthening is proportional to tension (at least in the case of not too large length changes), and therefore,

$$f \simeq \phi \frac{L - L_0}{N}, \qquad |L - L_0| \ll L_0,$$

where $L_0 = N\ell_0$ is the length at rest.

1. Derive the analog of the Gibbs-Duhem equation.
2. Show that ϕ is proportional to the temperature T:

   $\phi = bT$.

3. Show that the internal energy E depends only on T, and that the entropy S is the sum of a term depending only on T and one depending only on L.
4. Show how to obtain the fundamental equation in both the entropy and the energy schemes.

Recommended Reading

The preceding discussion was inspired by H. B. Callen's *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed., New York: J. Wiley & Sons, 1985. Callen's ideas are in their turn rooted in those of L. Tisza: see L. Tisza, *Generalized Thermodynamics*, Cambridge, MA: The MIT Press, 1966. More "traditional" discussions follow a different path: their goal is to construct the axiomatic structure starting from the phenomenology. An excellent presentation can be found in A. B. Pippard, *The Elements of Classical Thermodynamics*, Cambridge, UK: Cambridge University Press, 1957.

3 The Fundamental Postulate

When he woke up, the dinosaur was still there.
—Augusto Monterroso

The aim of statistical mechanics is to describe the thermodynamic properties of systems composed of a large number of particles. The time scales of the evolution of these systems are microscopic, and for this reason, measuring mechanical quantities in an experiment of reasonable length is equivalent to measuring the averages of these quantities during the corresponding time interval. In this manner, it is possible to define the thermodynamic observables as time averages of mechanical quantities. For example, the pressure applied by a fluid on a wall can be expressed (by analogy with what we saw in chapter 1) as the time average of the force applied on the wall, divided by the surface area of the wall itself.

We have seen, however, that entropy is not a mechanical quantity, in the sense we just expressed. The key to the relation between mechanical and thermodynamic properties is the microscopic expression of entropy. This expression was found by Ludwig Boltzmann, and it plays such an important role that we have decided to call it the **fundamental postulate** of statistical mechanics. Since we are dealing with a postulate, it cannot be proven—we can, however, make it plausible by showing that the properties of entropy that derive from it agree with the properties that we postulated in thermodynamics.

In order to make the fundamental postulate more explicit, we must resort to an abstract representation of the dynamics of a system with many particles—this representation uses the concepts of phase space and of a system's representative point.

3.1 Phase Space

Let us consider a simple fluid represented as a classical system composed of N identical particles, each of mass m, interacting with one another and with the walls of the container. Let us suppose that the system is isolated so that there are no interactions between it and the external environment. At each instant, the state of the system can be described by giving the position $\boldsymbol{r}_i = (r_{x,i}, r_{y,i}, r_{z,i})$ and the momentum $\boldsymbol{p}_i = (p_{x,i}, p_{y,i}, p_{z,i})$ of each particle i: $6N$ scalar variables in total. These identify a point x, the system's **representative point**,

within a $6N$-dimensional space, called the **phase space**:

$$x = \left(\boldsymbol{p}^N, \boldsymbol{r}^N\right), \tag{3.1}$$

where

$$\boldsymbol{p}^N = \left(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N\right), \qquad \boldsymbol{r}^N = \left(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\right). \tag{3.2}$$

We shall say that the representative point identifies a **microscopic state** (or **microstate**) of the system. The evolution of the system is represented by the motion of the representative point within this space and is described by Hamilton's system of $6N$ differential equations of motion. Given, for example, a system of Hamiltonian $\mathcal{H}$ of the form

$$\mathcal{H}(x) = \sum_{i=1}^{N} \frac{p_i^2}{2m} + U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N), \tag{3.3}$$

its corresponding Hamilton's equations are given by

$$\frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{r}_i}, \qquad \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}_i}, \qquad i = 1, \ldots, N. \tag{3.4}$$

With the passage of time, the system's representative point moves along a trajectory $x = x(t; x_0)$, identified by the initial condition $x(0; x_0) = x^0 = (\boldsymbol{p}^0, \boldsymbol{r}^0)$.

Generally speaking, this trajectory is extremely complicated. The only thing we can state a priori is that, if the system is isolated, it lies entirely on the surface defined by the equation

$$\mathcal{H}(x) = E = \text{const.} \tag{3.5}$$

The constant value E of the Hamiltonian can be identified as the system's internal energy.

Strictly speaking, for a system of $6N$ equations of motion, there exist $6N - 1$ integrals of motion, which define a line in phase space along which the system must move. To a very large extent, however, these integrals cannot be measured or theoretically predicted, because a small disturbance to the system causes their values to vary greatly. In a system composed of many particles, a small (experimentally imperceptible) shift of each particle corresponds to a large shift of the system's representative point in phase space. It is obvious that quantities that change a lot due to a shift of this kind cannot be relevant to a thermodynamic description of the system. Therefore, the properties that are dealt with by thermodynamics and statistical mechanics must vary only slightly when the system's representative point is changed slightly. We make this condition more explicit later.

In order for our argument to be consistent with thermodynamics, it is necessary that given a certain (not very large) number of these quantities, one be able to predict their value (and the value of other analogous quantities) after thermodynamic transformations have occurred, at least for sufficiently large systems. How these quantities can be characterized is a well-defined problem, but one that has not yet been satisfactorily resolved in the general case. In order to obtain a microscopic equivalent of the postulates of thermodynamics, however, we can assume that these quantities are the thermodynamic observables, only

this time considered as functions of the *microscopic state* of the system, rather than of its thermodynamic state. We shall therefore formulate the following hypothesis:

The region in phase space in which the system is evolving can be completely identified, for thermodynamic purposes, by the values of certain observable quantities.

In the case of the simple fluid we considered, this region can be identified by the condition that all particles be situated inside a region of volume V, which represents the inside of the container, and by the value E of the internal energy. If the thermodynamics of simple fluids that we discussed in chapter 2 is valid, it is impossible to identify other physically important quantities that vary slowly in the phase space and that restrict this region in a significant manner.

3.2 Observables

Let us call **observables** in statistical mechanics those functions, defined on the phase space, that vary smoothly enough when the microstate varies. For example, the total kinetic energy of a mechanical system defined by

$$K = \sum_{i=1}^{N} \frac{p_i^2}{2m} \tag{3.6}$$

is an observable—a variation in momentum of a particle causes a relative increase in K on the order of N^{-1} (which is negligible when N is large).

An example of a *nonobservable* function is the following:

$$\Theta = \prod_{i=1}^{N} \theta_{\mathrm{H}} \left(r_0^2 - r_i^2 \right), \tag{3.7}$$

where $\theta_{\mathrm{H}}(x)$ is **Heaviside's theta function**, defined by

$$\theta_{\mathrm{H}}(x) = \begin{cases} 1, & \text{if } x > 0; \\ 1/2, & \text{if } x = 0; \\ 0, & \text{if } x < 0. \end{cases} \tag{3.8}$$

Then Θ is equal to 1 if all the particles are within a sphere of radius r_0 around the origin and is otherwise equal to 0. A shift applied to just one particle is sufficient to change the value of this function by a finite quantity.

As we saw previously, the internal energy corresponds to the observable $\mathcal{H}$ (the Hamiltonian). The number of particles can be expressed as an observable as follows:

$$N = \int \mathrm{d}\boldsymbol{r} \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i). \tag{3.9}$$

We are thus led to the following hypothesis:

All the extensive thermodynamic variables appearing as independent variables in the fundamental equation are observables.

Since the fundamental equation is the expression of entropy as a function of extensive variables, what we are missing is a mechanical interpretation of entropy. The fundamental postulate provides such an interpretation.

In order to take advantage of the extensivity, we shall consider our systems in the **thermodynamic limit**, in which the size of the system, as given, for example, by the number of particles N, grows without limit, but the ratios of the extensive variables, that is, the densities such as E/N, V/N, etc., are kept constant.

3.3 The Fundamental Postulate: Entropy as Phase-Space Volume

The fundamental postulate of statistical mechanics expresses entropy as a function of the accessible volume in phase space—in other words, of the volume of the phase space in which the thermodynamic observables have values compatible with a specific thermodynamic state.

Let us assume that the thermodynamic state is determined by the extensive variables $(X_0, \ldots, X_r)$. Each of these variables—X_i, for example—is represented (in our hypotheses) by a function $X_i(x)$ defined over the space of microscopic states. The region of phase space that is accessible is defined by the values assumed by these variables which we will denote by $(X_0, \ldots, X_r)$. In effect, this region is defined by the values $(X_0, \ldots, X_r)$, up to errors ΔX_i, $i = 1, \ldots, r$, which can be assumed to be small but macroscopic. The region of phase space in which these observables have those specific values (within the errors) will be denoted by Γ, and its volume by $|\Gamma|$. Then the fundamental postulate states that

$$S = k_{\mathrm{B}} \ln |\Gamma| \,. \tag{3.10}$$

In this formula,

- S is the thermodynamic entropy (expressed as a function of extensive variables),
- k_{B} is Boltzmann's constant ($\simeq 1.381 \cdot 10^{-23} \mathrm{JK}^{-1}$), and
- the equals sign must be understood as implying "except for subextensive terms, which become negligible in the thermodynamic limit."

Since it is a postulate, this relation cannot be "proven"—however, one can attempt to justify it by showing that $\ln |\Gamma|$ has the properties of thermodynamic entropy.

We denote by $\mathrm{d}x$ the volume element of the phase space:

$$\mathrm{d}x = \prod_{i=1}^{N} \left(\mathrm{d}\boldsymbol{p}_i \, \mathrm{d}\boldsymbol{r}_i\right) = \mathrm{d}\boldsymbol{p}^N \, \mathrm{d}\boldsymbol{r}^N, \tag{3.11}$$

where we have introduced the shorthands

$$\mathrm{d}\boldsymbol{p}^N = \prod_{i=1}^{N} \mathrm{d}\boldsymbol{p}_i, \qquad \mathrm{d}\boldsymbol{r}^N = \prod_{i=1}^{N} \mathrm{d}\boldsymbol{r}_i. \tag{3.12}$$

Then the preceding discussion allows one to give the following expression to thermodynamic entropy, as a function of the extensive variables $(X_0, \ldots, X_r)$:

$$\begin{aligned} S(X_0, \ldots, X_r) &= k_{\mathrm{B}} \ln \int_{\Gamma} \mathrm{d}x \\ &= k_{\mathrm{B}} \ln \int \mathrm{d}x \prod_{i=0}^{r} \left[\theta_{\mathrm{H}}\left(X_i(x) - X_i - \Delta X_i\right) \theta_{\mathrm{H}}\left(X_i - X_i(x)\right)\right]. \end{aligned} \tag{3.13}$$

The product of the theta functions assumes the value 1 if the microscopic state x is such that the values of the observables belong to the interval $X_i - \Delta X_i \leq X_i(x) \leq X_i$, $i = 0, \ldots, r$, and is otherwise equal to zero. Thus the integral is equal to the volume of the phase space region Γ where these inequalities are satisfied.

3.3.1 Other Expressions of the Entropy

If we assume that the fundamental postulate is valid, as a corollary we can conclude that the quantity that appears on the left-hand side of eq. (3.13) is extensive. As a consequence, it is possible to obtain other expressions of the same quantity, ones that differ from it for terms that grow less rapidly than S with the system size. Since in what follows we will make use of these expressions in an interchangeable manner, it is useful to discuss them explicitly here. For simplicity we consider a situation in which all the extensive thermodynamic variables $(X_1, \ldots, X_r)$ have a well-defined value, and we explicitly consider entropy solely as a function of internal energy E.

In this situation, eq. (3.13) takes the form

$$S(E) = k_{\mathrm{B}} \ln \int \mathrm{d}x\, \theta_{\mathrm{H}}(\mathcal{H}(x) - E - \Delta E)\, \theta_{\mathrm{H}}(E - \mathcal{H}(x)). \tag{3.14}$$

A fairly surprising first result is that if the energy has a lower bound (we may assume, for example, that it cannot be negative), one obtains the same value for entropy by extending the integral to the entire region $\mathcal{H}(x) \leq E$:

$$S(E) = k_{\mathrm{B}} \ln \int \mathrm{d}x\, \theta_{\mathrm{H}}(E - \mathcal{H}(x)) + \text{subextensive terms}. \tag{3.15}$$

Indeed, as we know, $S(E)$ is a monotonically increasing function of E, and it is extensive. The accessible phase space volume $\left|\Gamma(E)\right| = \mathrm{e}^{S(E)/k_{\mathrm{B}}}$ increases with E, and it increases all the more rapidly the larger the size of the system being considered. As a consequence, the greater part of the accessible phase space volume corresponds to the largest values of E. In effect, given two values of internal energy, E and E', the ratio of the corresponding phase space volumes is given by

$$\frac{\left|\Gamma(E)\right|}{\left|\Gamma(E')\right|} = \mathrm{e}^{(S(E) - S(E'))/k_{\mathrm{B}}}, \tag{3.16}$$

and if $E < E'$, this ratio is exponentially small in the system size. We have thus obtained expression (3.15).

Given the same hypotheses, we can conclude that the entropy can also be expressed as a function of the area of the surface defined by $\mathcal{H}(x) = E$:

$$S(E) = k_{\mathrm{B}} \ln \int \mathrm{d}x\, \delta\left(\mathcal{H}(x) - E\right) + \text{subextensive terms}, \tag{3.17}$$

where $\delta(x)$ is the Dirac delta "function" (cf. sec. A.2). In effect, let us temporarily denote the quantity defined on the right-hand side of eq. (3.17) by $\mathfrak{S}(E)$. By substituting this expression in (3.13), we obtain

$$S(E) = k_{\mathrm{B}} \ln \int_{E-\Delta E}^{E} \mathrm{d}E'\, \mathrm{e}^{\mathfrak{S}(E')/k_{\mathrm{B}}}, \tag{3.18}$$

If ΔE is sufficiently small but finite, the integrand varies slowly in the interval $E - \Delta E \leq E' \leq E$, and one therefore arrives at

$$S(E) \simeq k_{\mathrm{B}} \ln\left(\mathrm{e}^{\mathfrak{S}(E)/k_{\mathrm{B}}}\, \Delta E\right) = \mathfrak{S}(E) + k_{\mathrm{B}} \ln \Delta E. \tag{3.19}$$

The last term is obviously negligible in the thermodynamic limit, and therefore $S(E) = \mathfrak{S}(E)$. By using this result, we can obtain a more explicit derivation of (3.15).

3.4 Liouville's Theorem

In the fundamental postulate, entropy is expressed in terms of the accessible volume in *phase space*. One may ask why phase space plays this privileged role, rather than other possible spaces that also describe the microstates of a mechanical system. Let us consider for example a system composed of identical particles of mass m. In this case, the momentum $\boldsymbol{p}$ of a particle is simply proportional to its velocity $\boldsymbol{v}$, and it is a matter of taste to describe the system in the phase space, that is, by the collection $(\boldsymbol{r}, \boldsymbol{p})$ of the position and momentum vectors of the particles or by the space of position and velocity vectors. But consider instead a mixture of particles of two different masses—say, m and m'. When particles collide, they exchange momentum in the center-of-mass frame. If they have the same mass, this is tantamount to a velocity exchange. But this is not true when the particles have different masses. Indeed, we can show that the phase space of positions and momenta possesses a very special property with respect to time evolution. This property is expressed by **Liouville's theorem**.

Let us consider a mechanical system described in the phase space $x = \left(\boldsymbol{p}^N, \boldsymbol{r}^N\right)$, where $\boldsymbol{p}^N = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N)$ and $\boldsymbol{r}^N = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$. Let the Hamiltonian of the system be $\mathcal{H}\left(\boldsymbol{p}^N, \boldsymbol{r}^N, t\right)$, which may depend on the time t in an arbitrary way. Then the representative point of the system evolves according to the canonical equations of motion (3.4). In this way, to each point x in phase space is assigned a velocity vector $b(x, t)$ in the same space such that if the representative point of the system lies at x at a given time t, it will be at $x + b(x, t)\mathrm{d}t$ at time $t + \mathrm{d}t$, where $\mathrm{d}t$ is infinitesimal. The vector $b = b(x, t)$ has $6N$ coordinates, given by the right-hand side of Hamilton's equations:

$$b_{r_{\alpha,i}} = \frac{\partial \mathcal{H}}{\partial p_{\alpha,i}}, \qquad b_{p_{\alpha,i}} = -\frac{\partial \mathcal{H}}{\partial r_{\alpha,i}}, \qquad \alpha \in \{x, y, z\}, \quad i = 1, \ldots, N. \tag{3.20}$$

Consider a region $\Gamma(0)$ of phase space. Given the time t, consider the region $\Gamma(t)$ spanned by $x(t; x^0)$ as x^0 spans $\Gamma(0)$. We can call $\Gamma(t)$ the **evolute** at time t of the region $\Gamma(0)$, since it is the locus of all states in which the system can be found at time t if its state at time 0 had been in $\Gamma(0)$. It is clear that the location and shape of $\Gamma(t)$ changes with the time t.

Liouville's theorem states that the volume $|\Gamma(t)|$ of $\Gamma(t)$ does not change with t, and that therefore it remains equal to the volume of $\Gamma(0)$:

$$\frac{\mathrm{d}\,|\Gamma(t)|}{\mathrm{d}t} = 0. \tag{3.21}$$

This property holds for all systems whose evolution is described by Hamilton's equations of motion, even if the Hamiltonian does not possess the form (3.3) and/or depends explicitly on time.

The volume of a region Γ of phase space is given by

$$|\Gamma| = \int_\Gamma \mathrm{d}x = \int_\Gamma \mathrm{d}\boldsymbol{p}^N\, \mathrm{d}\boldsymbol{r}^N. \tag{3.22}$$

It is then clear that Liouville's theorem will be proved if we are able to prove it for an infinitesimal region $\mathrm{d}\Gamma(0)$, such that $x \in \mathrm{d}\Gamma$ if it satisfies the inequalities

$$p^0_{\alpha,i} \le p_{\alpha,i} \le p^0_{\alpha,i} + \mathrm{d}p_{\alpha,i}, \qquad r^0_{\alpha,i} \le r_{\alpha,i} \le r^0_{\alpha,i} + \mathrm{d}r_{\alpha,i}, \tag{3.23}$$

where $\alpha \in \{x, y, z\}$ and $i = 1, \ldots, N$. Indeed, regions of arbitrary sizes can always be considered to be made of the union of a large number of infinitesimal ones. We shall actually prove a stronger statement, namely that the projection of $\mathrm{d}\Gamma(t)$ in each of the two-dimensional subspaces spanned by the axes $p_{\alpha,i}$ and $r_{\alpha,i}$, with $\alpha \in \{x, y, z\}$ and $i = 1, \ldots, N$, remains constant in time as the system evolves. It is sufficient to show that this property holds for a time interval of infinitesimal duration $\mathrm{d}t$, since it is clear that if the volume of the evolute remains constant for a small time change, iterating the statement will prove that it remains constant forever.

Let us thus consider one such subspace, and denote by (r, p) its corresponding coodinates. The region $\mathrm{d}\Gamma(0)$ is defined by the inequalities

$$p^0 \le p \le p^0 + \mathrm{d}p, \qquad r^0 \le r \le r^0 + \mathrm{d}r, \tag{3.24}$$

and is represented in fig. 3.1 by the rectangle of corners A, B, C, and D. After a short time interval of duration $\mathrm{d}t$ the region will be approximately a parallelogram of corners A′, B′, C′, and D′. While the area of $\mathrm{d}\Gamma(0)$ is given by $\mathrm{d}r\,\mathrm{d}p$, the area of this parallelogram is given by the vector product of $\overrightarrow{\mathrm{A'B'}}$ with $\overrightarrow{\mathrm{A'D'}}$. We now have, by applying the equations of motion, and expanding to first order in $\mathrm{d}r$ and $\mathrm{d}p$,

$$\begin{aligned}
\overrightarrow{\mathrm{AA'}} &= \left(-\frac{\partial \mathcal{H}}{\partial r}\,\mathrm{d}t, \frac{\partial \mathcal{H}}{\partial p}\,\mathrm{d}t \right); \\
\overrightarrow{\mathrm{BB'}} &= \left(-\frac{\partial \mathcal{H}}{\partial r}\,\mathrm{d}t - \frac{\partial^2 \mathcal{H}}{\partial p\,\partial r}\mathrm{d}p\,\mathrm{d}t, \frac{\partial \mathcal{H}}{\partial p}\,\mathrm{d}t + \frac{\partial^2 \mathcal{H}}{\partial p^2}\mathrm{d}p\,\mathrm{d}t \right); \\
\overrightarrow{\mathrm{DD'}} &= \left(-\frac{\partial \mathcal{H}}{\partial r}\,\mathrm{d}t - \frac{\partial^2 \mathcal{H}}{\partial r^2}\mathrm{d}r\,\mathrm{d}t, \frac{\partial \mathcal{H}}{\partial p}\,\mathrm{d}t + \frac{\partial^2 \mathcal{H}}{\partial r\,\partial p}\mathrm{d}r\,\mathrm{d}t \right).
\end{aligned} \tag{3.25}$$

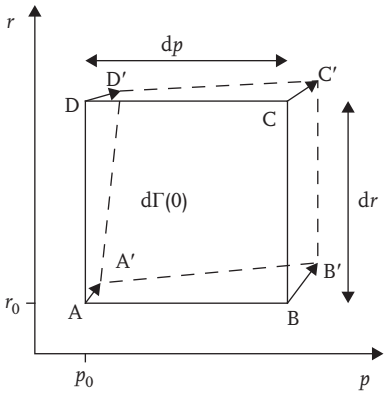


FIGURE 3.1. Derivation of the Liouville theorem. The region $d\Gamma(0)$, defined by the inequalities (3.24), is represented by the rectangle with full lines of area $dp\,dr$. Its evolute after a short time dt is represented by the quadrilateral with dashed lines, whose area is approximately given by the vector product of $\overrightarrow{A'B'}$ with $\overrightarrow{A'D'}$.

Here all derivatives are supposed to be evaluated in the point (p^0, r^0). Taking into account that $\overrightarrow{AB} = (dp, 0)$ and $\overrightarrow{AD} = (0, dr)$, we have

$$\begin{aligned}\overrightarrow{A'B'} &= \overrightarrow{AB} + \overrightarrow{BB'} - \overrightarrow{AA'} = \left(dp - \frac{\partial^2\mathcal{H}}{\partial r\,\partial p}dp\,dt,\ \frac{\partial^2\mathcal{H}}{\partial p^2}dp\,dt\right);\\ \overrightarrow{A'D'} &= \overrightarrow{AD} + \overrightarrow{DD'} - \overrightarrow{AA'} = \left(-\frac{\partial^2\mathcal{H}}{\partial r^2}dr\,dt,\ dr + \frac{\partial^2\mathcal{H}}{\partial p\,\partial r}dr\,dt\right).\end{aligned} \tag{3.26}$$

We thus obtain

$$\begin{aligned}\overrightarrow{A'B'} \times \overrightarrow{A'D'} &= \begin{vmatrix} dp - (\partial^2\mathcal{H}/\partial r\,\partial p)dp\,dt, & \partial^2\mathcal{H}/\partial p^2\,dp\,dt \\ -(\partial^2\mathcal{H}/\partial r^2)dr\,dt, & dr + \partial^2\mathcal{H}/\partial p\,\partial r\,dr\,dt \end{vmatrix}\\ &= dp\,dr\left[1 + O\left(dt^2\right)\right],\end{aligned} \tag{3.27}$$

since the terms of order dt in the vector product cancel out by the symmetry of the mixed derivatives of $\mathcal{H}$. This concludes the proof.

The fact that the canonical equations conserve the volume of regions in phase space is the special property of phase space we were looking for.

We can prove Liouville's theorem directly in the phase space by a different route. We can envisage the evolution $x(t)$ of the representative point of our system as a flow of a fluid in phase space, whose local velocity is given by eq. (3.20). Let us now consider a closed surface in this space. It is possible to generalize Gauss's theorem to such a space, to show that the flux of the velocity field across the surface is equal to the integral of the divergence of the field evaluated in the interior of the surface. Keeping this in mind, let us evaluate the divergence of the vector field $b(x, t)$:

$$\operatorname{div} b = \sum_{\alpha\in\{x,y,z\}}\sum_{i=1}^{N}\left[\frac{\partial}{\partial p_{\alpha,i}}\left(-\frac{\partial\mathcal{H}}{\partial r_{\alpha,i}}\right) + \frac{\partial}{\partial r_{\alpha,i}}\frac{\partial\mathcal{H}}{\partial p_{\alpha,i}}\right] = 0, \tag{3.28}$$

by the symmetry of the mixed derivatives. It follows that the flux of $b(x, t)$ across any closed surface vanishes. Given such a surface, the flux of b times dt measures the net volume of the fluid that crosses the surface in a short time interval of duration dt. But this quantity

is equal, up to the sign, to the difference in volume between the region contained by the surface and its evolute after a time $\mathrm{d}t$. Since $\operatorname{div} b = 0$, the flux vanishes for any closed surface, and therefore the volume of the evolute remains equal to the volume of the original region.

We can now imagine a great number of identical systems, whose representative points are initially distributed in phase space with a density $\rho(x^0, t=0)$ at time $t=0$. These points move according to Hamilton's equation of motion. We can picture the collective motion of these points as a fluid flow in phase space. Due to Liouville's theorem, this motion has special properties. For instance, given an infinitesimal region $\mathrm{d}\Gamma(0)$ around x^0 the number $\mathrm{d}N$ of points contained in it is given by $\mathrm{d}N = \rho(x^0, t=0)\ \left|\mathrm{d}\Gamma(0)\right|$. After a time t, the number of points contained in the evolute $\mathrm{d}\Gamma(t)$ of $\mathrm{d}\Gamma(0)$ is obviously still equal to $\mathrm{d}N$, and is given by $\rho(x(t; x^0), t)\ \left|\mathrm{d}\Gamma(t)\right|$. But, since $\left|\mathrm{d}\Gamma(t)\right| = \left|\mathrm{d}\Gamma(0)\right|$ by Liouville's theorem, we have

$$\rho(x(t; x^0), t) = \rho(x^0, t=0), \qquad \forall t. \tag{3.29}$$

In other words, the total derivative with respect to time of the density of representative points in phase space vanishes:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\partial\rho}{\partial t} + \sum_{i=1}^{N}\left(\frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t}\frac{\partial\rho}{\partial\boldsymbol{p}_i} + \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t}\frac{\partial\rho}{\partial\boldsymbol{r}_i}\right) = 0. \tag{3.30}$$

We can write this relation more explicitly by substituting the equations of motion, and we obtain

$$\frac{\partial\rho}{\partial t} = \sum_{i=1}^{N}\left(\frac{\partial\mathcal{H}}{\partial\boldsymbol{r}_i}\frac{\partial\rho}{\partial\boldsymbol{p}_i} - \frac{\partial\mathcal{H}}{\partial\boldsymbol{p}_i}\frac{\partial\rho}{\partial\boldsymbol{r}_i}\right) = -\left[\mathcal{H}, \rho\right]_{\mathrm{PB}}, \tag{3.31}$$

where $[A, B]_{\mathrm{PB}}$ denotes the **Poisson bracket** of A and B:

$$[A, B]_{\mathrm{PB}} = \sum_{i=1}^{N}\left(\frac{\partial A}{\partial\boldsymbol{p}_i}\frac{\partial B}{\partial\boldsymbol{r}_i} - \frac{\partial A}{\partial\boldsymbol{r}_i}\frac{\partial B}{\partial\boldsymbol{p}_i}\right). \tag{3.32}$$

Equation (3.31) is called the **Liouville equation**. Its right-hand side is a linear operator acting on the phase distribution ρ, usually called the **Liouvillian**. Denoting it by $-\mathcal{L}$ we can rewrite Liouville's equation in the concise form

$$\frac{\partial\rho}{\partial t} = -\mathcal{L}\,\rho. \tag{3.33}$$

Let $A(x)$ be an observable, that is, a function defined on the phase space of the system. Let us evaluate the change in its instantaneous value as the system evolves according to the canonical equations of motion. We set

$$A(t) = A(x(t; x^0)), \tag{3.34}$$

and we have

$$\begin{aligned}\frac{\mathrm{d}A}{\mathrm{d}t} &= \sum_{\alpha\in\{x,y,z\}}\sum_{i=1}^{N}\left(b_{r\alpha,i}\frac{\partial A}{\partial r_{\alpha,i}} + b_{p\alpha,i}\frac{\partial A}{\partial p_{\alpha,i}}\right)\\ &= \sum_{\alpha\in\{x,y,z\}}\sum_{i=1}^{N}\left(\frac{\partial \mathcal{H}}{\partial p_{\alpha,i}}\frac{\partial A}{\partial r_{\alpha,i}} - \frac{\partial \mathcal{H}}{\partial r_{\alpha,i}}\frac{\partial A}{\partial p_{\alpha,i}}\right) = [\mathcal{H}, A]_{\mathrm{PB}}\,.\end{aligned} \tag{3.35}$$

Thus, the derivative of A with respect to time is expressed in terms of its Poisson bracket with the Hamiltonian. Notice, however, the change in sign with respect to Liouville's equation (3.31). In particular, the constants of the motion are characterized by the fact that their Poisson bracket with the Hamiltonian vanishes.

Let us now assume that $\rho(x,t)$ is such that, at $t=0$, it depends only on the values of a set $(X_1,\ldots,X_r)$ of constants of the motion, that is,

$$\rho(x,0) = \Phi\left(X_1(x)\ \ldots, X_r(x)\right), \tag{3.36}$$

where, for $k=1,\ldots,r$, one has

$$\frac{\mathrm{d}X_k}{\mathrm{d}t} = [\mathcal{H}, X_k]_{\mathrm{PB}} = 0. \tag{3.37}$$

We then have, at $t=0$,

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\sum_{\alpha\in\{x,y,z\}}\sum_{i=1}^{N}\left(\frac{\partial \mathcal{H}}{\partial p_{\alpha,i}}\frac{\partial \rho}{\partial r_{\alpha,i}} - \frac{\partial \mathcal{H}}{\partial r_{\alpha,i}}\frac{\partial \rho}{\partial p_{\alpha,i}}\right)\\ &= -\sum_{k=1}^{r}\sum_{\alpha\in\{x,y,z\}}\sum_{i=1}^{N}\left(\frac{\partial \mathcal{H}}{\partial p_{\alpha,i}}\frac{\partial X_k}{\partial r_{\alpha,i}} - \frac{\partial \mathcal{H}}{\partial r_{\alpha,i}}\frac{\partial X_k}{\partial p_{\alpha,i}}\right)\frac{\partial \Phi}{\partial X_k}\\ &= \sum_{k=1}^{r}[\mathcal{H}, X_k]_{\mathrm{PB}}\,\frac{\partial \Phi}{\partial X_k} = 0.\end{aligned} \tag{3.38}$$

Therefore, $\rho(x,t)$ does not change with time at $t=0$, nor changes afterward. This provides us with a recipe to obtain distributions in phase space that do not change with time—it is sufficient that they depend on x only via some constants of the motion X_k.

3.5 Quantum States

In classical thermodynamics, entropy is defined up to an additive constant. To remove this ambiguity, one adds a third postulate to the principles discussed in the preceding chapter, a postulate called **Nernst's postulate** or the **third law of thermodynamics**, which stipulates the following property:

The entropy of any system approaches zero for $T \to 0$.

According to this postulate, each isotherm transformation at $T=0$ is also an adiabatic transformation. (From this, it follows that one cannot reach absolute zero.)

On the other hand, from the expression $S = k_B \ln |\Gamma|$ of the fundamental postulate, it is clear that if we change the units of measurement of the phase space, we change the value of the entropy by an additive constant. As a consequence, if Nernst's postulate univocally sets the value of entropy (except in cases involving subextensive terms), it must also implicitly fix the unit of measurement for the volume of the phase space.

The value of this unit is obtained from considerations that involve quantum mechanics and that I attempt to describe briefly in this section. We have to keep in mind that we are still far from having a satisfying set of postulates for the foundation of quantum statistical mechanics. The informal remarks that follow have only a heuristic value.

A microscopic state in quantum mechanics is described by a vector (normalized and defined up to a phase) $|\Psi\rangle$ that belongs to a Hilbert space $\mathbb{H}$. It evolves according to the time-dependent Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} |\Psi\rangle = \widehat{\mathcal{H}} |\Psi\rangle , \tag{3.39}$$

where $\widehat{\mathcal{H}}$ is the Hamiltonian operator and $\hbar = h/2\pi$, where h is Planck's constant ($h \simeq 6.626 \cdot 10^{-34}$ Js). (We always denote quantum operators with a $\widehat{\ }$.) The quantum observables A are represented by Hermitian operators $\widehat{A}$ defined on $\mathbb{H}$, whose eigenvalues correspond to the possible results of the measurement of A.

More particularly, if the operator $\widehat{A}$ commutes with the Hamiltonian $\widehat{\mathcal{H}}$, it is a constant of motion, and its average value does not change over time—in addition, it is possible to identify some common eigenstates $|E, A\rangle$ of both the Hamiltonian and $\widehat{A}$ in which both the energy E and the observable A have well-defined values. If a system is prepared in one of these eigenstates, the evolution of the system (dictated by the Schrödinger equation (3.39)) keeps it in the same state.

Since the observables X are extensive, their commutators are also extensive, which implies that the quantum uncertaintes related to the simultaneous mesurement of different observables are subextensive. It is therefore in principle possible to identify a set of extensive operators $(\widehat{X}_i)$ that differ from the original ones by subextensive amounts, and that commute among themselves and with the Hamiltonian. We shall consider these as the observables from now on.

Let us therefore consider the set $\mathfrak{S}$ of all microscopic states in which the thermodynamic variables have a well-defined value, equal to $(X_0, X_1, \ldots, X_r)$ up to the small uncertanties ΔX_i, $i = 0, \ldots, r$. These states are the common eigenstates of the Hamiltonian and of the operators $(\widehat{X}_0, \widehat{X}_1, \ldots, \widehat{X}_r)$, representing the thermodynamic observables, and belonging to eigenvalues Ξ_i, $i = 0, \ldots, r$, that satisfy the inequalities $X_i \leq \Xi_i \leq X_i + \Delta X_i$, $i = 0, \ldots, r$. The space spanned by $\mathfrak{S}$ states is the quantum analog of the accessible space Γ; and its dimensionality $\mathcal{N}$, that is, the number of linearly independent states that generate $\mathfrak{S}$, is the analog of the volume $|\Gamma|$ of the accessible phase space.

The fundamental postulate, in its quantum version, then stipulates that

$$S = k_B \ln \mathcal{N}. \tag{3.40}$$

In this formulation, the arbitrariness associated with the unit of measurement of the phase space volume has disappeared. Moreover, the Nernst postulate has a clear interpretation:

when the temperature tends to zero, only those states that correspond to the lowest energy values become accessible. The number of these states is finite, or grows more slowly than an exponential when the system's site increases, so that the extensive part of the entropy vanishes.[1]

We can now see how in the semiclassical limit, the number $\mathcal{N}$ of quantum states is proportional to the volume $|\Gamma|$ of the phase space, and we can calculate the relevant proportionality constant. In effect, Heisenberg's indeterminacy principle stipulates that for each pair (q, p) formed by a generalized coordinate q and its conjugate momentum p, the quantum uncertainties satisfy

$$\Delta p\, \Delta q \geq \mathrm{O}(h). \tag{3.41}$$

Each quantum state therefore "occupies" a region in phase space whose projection in the (q, p) plane has an area proportional to h. For a system with N degrees of freedom, the volume is therefore proportional to h^N, where N is the number of pairs of coordinates with their relevant conjugate momenta.

In order to determine the constant of proportionality (which must be universal), one can consider some simple systems. Let us take the harmonic oscillator, for example. The eigenstates of the Hamiltonian of a harmonic oscillator with angular frequency ω correspond to the values $E_n = \hbar\omega(n + 1/2)$ of energy, where $n = 0, 1, \ldots$. We can evaluate the volume of phase space "occupied" by each eigenstate by calculating the volume included between the curves $E = n\hbar\omega$ and $E = (n + 1)\hbar\omega$. (Since we are in two dimensions, the "volume" is, obviously, an area.) We have

$$\mathcal{H}(p, q) = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2, \tag{3.42}$$

and therefore

$$\Delta\,|\Gamma| = \int_{n\hbar\omega}^{(n+1)\hbar\omega} \mathrm{d}E \int \mathrm{d}p\, \mathrm{d}q\, \delta\left(\mathcal{H}(p, q) - E\right). \tag{3.43}$$

We can now introduce the variables

$$x^2 = \frac{p^2}{m}, \qquad y^2 = m\omega^2 q^2, \tag{3.44}$$

and therefore the integral becomes

$$\begin{aligned} \Delta\,|\Gamma| &= \int_{n\hbar\omega}^{(n+1)\hbar\omega} \mathrm{d}E \int \frac{\mathrm{d}x\, \mathrm{d}y}{\omega}\, \delta\left(\frac{1}{2}(x^2 + y^2) - E\right) \\ &= \int_{2n\hbar\omega}^{2(n+1)\hbar\omega} \mathrm{d}r\, \frac{2\pi r}{\omega} = 2\pi\hbar = h. \end{aligned} \tag{3.45}$$

[1] One often comes across the statement that Nernst's postulate is a consequence of the fact that the ground state of a quantum system is unique. This is not correct, because the postulate concerns the behavior of the extensive part of $\ln \mathcal{N}$ for a system with a large N, in the limit $T \to 0$. (In other words, one first passes to the thermodynamic limit $N \to \infty$, and one then puts T to 0.) The uniqueness of the ground state is not sufficient to guarantee that entropy vanishes within this order of limits.

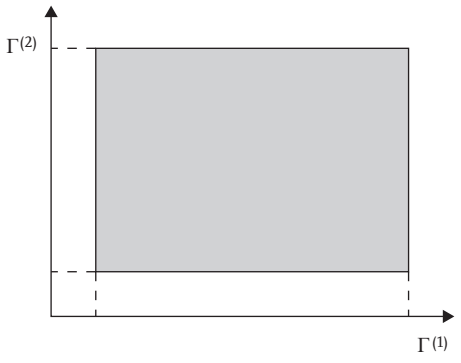


FIGURE 3.2. Schematic representation of the phase-space distribution $\Gamma^{(1)} \times \Gamma^{(2)}$ of two systems in contact.

The "volume" of each quantum state is therefore equal to h for a system with only one degree of freedom, as we had anticipated.

Exercise 3.1 (Particle in a potential well) Let us consider a particle of mass m, constrained to moving in one dimension along a segment of length n. In this case, the possible energy values are $E_n = \hbar^2\pi^2 n^2/L^2$, where $n = 1, 2, \ldots$. Prove that in this case, too, the phase space volume "occupied" by each quantum state is equal to h.

More generally, one will be able to obtain the correct factor for calculating the volume of quantum states by comparing the entropy calculated classically with the quantum entropy in the semiclassical limit. (Usually, at high temperatures both system behave "classically.") Since we do not intend to deal with the specific problems of quantum statistical mechanics in this text, we will not delve more deeply into this topic.

3.6 Systems in Contact

We will now discuss the extensivity of the entropy. Let us consider two systems, 1 and 2, such that the thermodynamic state of 1 corresponds to the region $\Gamma^{(1)}$ of phase space, and that of 2 to $\Gamma^{(2)}$. Then the Cartesian product $\Gamma^{(1\cup 2)} = \Gamma^{(1)} \times \Gamma^{(2)}$ corresponds to the region of phase space describing the state of the compound system $1 \cup 2$ (see figure 3.2). The volume of this region is equal to $\left|\Gamma^{(1\cup 2)}\right| = \left|\Gamma^{(1)}\right| \left|\Gamma^{(2)}\right|$.

In this manner, the fundamental postulate gives us

$$\begin{aligned} S &= k_B \ln \left|\Gamma^{(1\cup 2)}\right| = k_B \ln \left(\left|\Gamma^{(1)}\right| \left|\Gamma^{(2)}\right|\right) \\ &= k_B \ln \left|\Gamma^{(1)}\right| + k_B \ln \left|\Gamma^{(2)}\right| = S^{(1)} + S^{(2)}, \end{aligned} \tag{3.46}$$

which confirms the additivity of the entropy.

We emphasize that this result implicitly presupposes that no relation between the dynamic states of each of the two subsystems exists, so that the variables that define the overall state of the system are given simply by those that define the state of each subsystem taken separately. This condition of dynamic independence of the subsystems is in principle

an experimentally based hypothesis, essential for the validity of statistical mechanics, and can be considered as a supplement to the fundamental postulate.

In the last analysis, this hypothesis corresponds to the independence (in the sense given it by probability theory) of the observables that pertain to different systems in contact. A little later, we will see to what extent it may be possible to justify this hypothesis on the basis of the chaos theory of dynamic systems.

3.7 Variational Principle

The other fundamental property of entropy is the **variational principle**, according to which it reaches a maximum at thermodynamic equilibrium. Let us see how this principle is compatible with the interpretation of entropy in terms of the accessible volume in phase space.

Let us consider a system whose thermodynamic state is identified by the values of a certain collection $X = (X_i)$ of extensive variables. At thermodynamic equilibrium these observables will have the equilibrium values X_{eq}:

The variational principle for the entropy states that the equilibrium values X_{eq} correspond to the maximum value of the entropy.

For example, let us assume that our system is composed of two systems in contact, 1 and 2, that can exchange energy. The total internal energy, E_{tot}, has a fixed value: at equilibrium, the value $E^{(1)}$ of the internal energy of system 1 will also be fixed. This value is determined by the condition that the entropy reaches a maximum. For simplicity, we will assume that the internal energy of each system is determined up to a (macroscopic) uncertainty ΔE, and that the thermodynamic states correspond to discrete values of energy that differ by ΔE.

If we consider the two systems as isolated from each other, the accessible volume of the phase space is given by

$$|\Gamma_{\text{in}}| = \left|\Gamma_{\text{in}}^{(1)}\right| \left|\Gamma_{\text{in}}^{(2)}\right| . \tag{3.47}$$

If we allow for energy exchange between the two systems, the accessible volume becomes

$$|\Gamma| = \sum_{E^{(1)}} \left|\Gamma^{(1)}(E^{(1)})\right| \left|\Gamma^{(2)}(E_{\text{tot}} - E^{(1)})\right| . \tag{3.48}$$

The sum runs over all possible values of $E^{(1)}$. It is obvious that $|\Gamma| \geq |\Gamma_{\text{in}}|$, since the term that appears in eq. (3.47) also contributes to eq. (3.48). Therefore the removal of the constraints leads to an increase in the accessible volume of the phase space—in other words, of the entropy. We thus obtain the property that entropy cannot decrease spontaneously.

At equilibrium, however, the value of the internal energy of system 1 is well determined, while in eq. (3.48) all possible values appear. If we were to take eq. (3.48) at face value, the result would be disastrous, because the value of the system's entropy would come to depend on its history, and not only on the values of the observables. In effect, this

equation implies

$$|\Gamma| > \left|\Gamma^{(1)}(E^{(1)}_{\text{eq}})\right| \left|\Gamma^{(2)}(E^{(2)}_{\text{eq}})\right|, \tag{3.49}$$

instead of the equality we would expect from thermodynamics.

This problem is solved by considering that the accessible volume of the phase space corresponds *almost entirely* to the states in which $E^{(1)} = E^{(1)}_{\text{eq}}$, $E^{(2)} = E^{(2)}_{\text{eq}}$, so that the difference between the two expressions of $\ln|\Gamma|$ is negligible for large values of N. This remark is in turn based on the extensivity of the entropy.

Let us make the sum explicit, reading the fundamental postulate backward. We thus obtain

$$\begin{aligned} |\Gamma| &= \sum_{E^{(1)}} \left|\Gamma^{(1)}(E^{(1)})\right| \left|\Gamma^{(2)}(E_{\text{tot}} - E^{(1)})\right| \\ &= \sum_{E^1} \exp\left\{\frac{1}{k_{\text{B}}}\left[S^{(1)}(E^{(1)}) + S^{(2)}(E_{\text{tot}} - E^{(1)})\right]\right\}. \end{aligned} \tag{3.50}$$

We know that the total entropy—and therefore the summand—reaches its maximum for the equilibrium values of $E^{(1)}$ and $E^{(2)}$. Let us approximate the sum with an integral and expand the exponential's argument into a Taylor series around these values. The first-order term vanishes, and we thus obtain

$$\begin{aligned} |\Gamma| \simeq \int \frac{\mathrm{d}E^{(1)}}{\Delta E} \exp\Bigg\{\frac{1}{k_{\text{B}}}\Big[&S^{(1)}(E^{(1)}_{\text{eq}}) + S^{(2)}(E^2_{\text{eq}}) \\ &+ \frac{1}{2}\left(\left.\frac{\partial^2 S^{(1)}}{\partial {E^{(1)}}^2}\right|_{E^{(1)}_{\text{eq}}} + \left.\frac{\partial^2 S^{(2)}}{\partial {E^{(2)}}^2}\right|_{E^{(2)}_{\text{eq}}}\right)\left(E^{(1)} - E^{(1)}_{\text{eq}}\right)^2\Big]\Bigg\}. \end{aligned} \tag{3.51}$$

Due to the extensivity of the entropy and of the internal energy, the derivatives $\partial^2 S/\partial E^2$ are of order $1/N$, where N is a measure of the two subsystems' sizes, and therefore the contribution of the integral in $E^{(1)}$ is a factor of order $N^{1/2}$. This produces an additive contribution to the entropy on the order of $\ln N$, which is therefore negligible.

This reasoning gives us another important piece of information: that each system's internal energy values can differ from the equilibrium value of quantities of order $N^{1/2}$ and are therefore negligible with respect to the value of internal energy, which is of order N. This is a first step toward a theory of the fluctuations of thermodynamic quantities.

We therefore see that the a priori uncertainty ΔE over the value of internal energy is not relevant, at least so long as it is not much smaller than the thermodynamic uncertainty, which is of order $N^{1/2}$—in accordance with thermodynamics, we therefore arrive at the conclusion that the value of entropy does not depend on the uncertainty by which we know the values of the thermodynamic variables.

3.8 The Ideal Gas

Let us now consider the ideal monoatomic gas. In this case, the extensive variables to be considered are the number of particles N, the volume V, and the internal energy E.

The system is described by the Hamiltonian:

$$\mathcal{H} = \sum_{i=1}^{N} \left[\frac{p_i^2}{2m} + v(\boldsymbol{r}_i) \right], \tag{3.52}$$

where the **single-particle potential** $v(\boldsymbol{r})$ is given by

$$v(\boldsymbol{r}) = \begin{cases} 0, & \text{if } \boldsymbol{r} \text{ is inside the box of volume } V, \\ \infty, & \text{otherwise.} \end{cases} \tag{3.53}$$

The accessible volume of the phase space is given by

$$|\Gamma| = \frac{1}{N!} \frac{1}{h^{3N}} \int \mathrm{d}\boldsymbol{p}^N \mathrm{d}\boldsymbol{r}^N \, \delta \left(\sum_i \left(\frac{p_i^2}{2m} + v(\boldsymbol{r}_i) \right) - E \right), \tag{3.54}$$

where $\mathrm{d}\boldsymbol{p}^N = \prod_{i=1}^{N} \mathrm{d}\boldsymbol{p}_i$ and analogously for $\mathrm{d}\boldsymbol{r}^N$. In this expression, the delta function in the momenta imposes the internal energy value, and the h^{-3N} factor is used to normalize the number of states, as previously mentioned. The factor $1/N!$ takes into account the fact that, because of the indistinguishability of identical particles, states that differ only by a permutation of particle indices correspond to the same quantum state and do not need to be counted separately. It is worthy of note that Gibbs was led to introducing this factor on the basis of considerations of entropy's extensivity, well before the introduction of quantum mechanics. Thus this factor is aptly called the **Gibbs correction factor**.

The integral over the positions $(\boldsymbol{r}_i)$ extends over the box of volume V because of the form of the confining potential (3.53) and yields a factor V^N. We are thus led to

$$|\Gamma| = \frac{V^N}{N! h^{3N}} \int \prod_{i=1}^{N} \mathrm{d}\boldsymbol{p}_i \, 2m \, \delta \left(\sum_i p_i^2 - 2mE \right). \tag{3.55}$$

We can interpret the right-hand side integral as the surface area of a sphere of radius $R = (2mE)^{1/2}$ in a space with $3N$ dimensions. It is easy to see (see appendix E) that this area is equal to $(1/2)\pi^{3N/2} R^{3N-1} / \Gamma_{\mathrm{E}}(3N/2)$, where $\Gamma_{\mathrm{E}}(z)$ is Euler's gamma function. We finally obtain

$$|\Gamma| = \frac{V^N \sqrt{\pi} (2mE)^{(3N-1)/2}}{N! h^{3N} \Gamma_{\mathrm{E}}(3N/2)}, \tag{3.56}$$

up to factors that vary more slowly with N. Shifting to logarithms and taking into account, the asymptotic expression of Euler's gamma function (Stirling's formula; see appendix E) we obtain

$$S = N k_{\mathrm{B}} \ln \left\{ \left(\frac{\mathrm{e} V}{N} \right) \left[\frac{\mathrm{e} 2\pi m}{h^3} \left(\frac{2E}{3N} \right) \right]^{3/2} \right\}, \tag{3.57}$$

up to terms that are negligible for large values of N. We can see from this expression that the entropy thus obtained is actually extensive (while it would not have been so if the factor $1/N!$ in the definition of $|\Gamma|$ had been absent).

It is interesting to note that the expression of the entropy of the ideal monoatomic gas, including the quantum factors h, was obtained in the early twentieth century by O. Sackur and H. M. Tetrode [Sack11, Tetr14], before the formulation of quantum mechanics. This expression is therefore known as the **Sackur-Tetrode formula**.

Having thus arrived at the fundamental equation, we can deduce the equations of state by taking derivatives:

$$\frac{1}{T} = \frac{\partial S}{\partial E}\bigg)_{V,N} = \frac{3}{2}\frac{Nk_B}{E}; \tag{3.58}$$

$$\frac{p}{T} = \frac{\partial S}{\partial V}\bigg)_{E,N} = \frac{Nk_B}{V}. \tag{3.59}$$

These equations correspond to the usual equations for the ideal monoatomic gas—let us note that they imply that the entropy has the expression $S = S^{(1)}(E) + S^{(2)}(V)$, and, moreover, that the internal energy of the ideal gas depends only on the temperature $E = E(T)$.

The reasoning that led Gibbs to introduce the $1/N!$ factor in eq. (3.54) can be summarized as follows. Consider two containers of equal volume, each containing the same number N of particles at the same temperature. Assume first that the particles are different from each other in nature. Connect now the two containers, allowing the particles to move from one container to the next. When equilibrium is reached, the accessible volume for each particle will be doubled, and therefore the entropy of the system will have increased by $Nk_B \ln 2$. This result is independent of the nature of the gases, except that the gases that are mixed must be of different kinds. On the other hand, if we start with two containers holding particles of the same species, and let the particles mix, there should be no increase in entropy. This is known as the **Gibbs paradox**. It is then clear that introducing the $1/N!$ factor solves the problem since it follows from eq. (3.57) that the entropy S has the same value for the connected containers as for the sum of the two unconnected ones. An interesting discussion of this topic is found in [Jayn92].

3.9 The Probability Distribution

The power of statistical mechanics is highlighted by the possibility of providing predictions for observables that are not explicitly considered in thermodynamics.

Let us assume that the state of our system is identified by a collection $X = (X_i)$ of extensive variables, which is associated with a region Γ of volume $|\Gamma|$. We want to calculate the value taken by another observable, A, which we will assume is extensive. This observable is defined by a function, $A(x)$, defined over the phase space. According to our hypothesis, in the thermodynamic state specified by the (X_i), the observable A assumes a specific equilibrium value, which we denote by A_{eq}. If we impose on our system the constraint that the value of A is not necessarily equal to the equilibrium value A_{eq}, the fundamental hypotheses of thermodynamics tell us that the entropy can only decrease:

$$S(X; A) \leq S(X; A_{eq}). \tag{3.60}$$

On the other hand, the fundamental postulate of statistical mechanics relates the entropy $S(X;A)$ to the volume $\left|\Gamma(A)\right|$ of the accessible phase space region where the further constraint is imposed that the value of A is given:

$$S(X;A) = k_{\mathrm{B}} \ln \int_{\Gamma} \mathrm{d}x\, \delta\left(A(x) - A\right). \tag{3.61}$$

Since the equilibrium value of the observable A is equal to A_{eq}, however, one must get

$$S(X) = S(X;A_{\mathrm{eq}}) \geq S(X;A). \tag{3.62}$$

These relations are compatible only if essentially the entire volume of the accessible region Γ of phase space corresponds to the equilibrium value A_{eq} of the observable A.

One can convince oneself that this is how things stand by evaluating the ratio of the volumes of the regions being considered:

$$\begin{aligned} \frac{\left|\Gamma(A)\right|}{|\Gamma|} &= \frac{1}{|\Gamma|} \int_{\Gamma} \mathrm{d}x\, \delta\left(A(x) - A\right) \\ &= \exp\left\{\frac{1}{k_{\mathrm{B}}}\left[S(X;A) - S(X;A_{\mathrm{eq}})\right]\right\} \\ &\simeq \exp\left\{-\frac{1}{2k_{\mathrm{B}}}\left[\left.\frac{\partial^2 S}{\partial A^2}\right|_{A_{\mathrm{eq}}} \left(A - A_{\mathrm{eq}}\right)^2\right]\right\}. \end{aligned} \tag{3.63}$$

The first-order contribution vanishes because of the maximum entropy principle. The second-order derivative is negative and of order $1/N$; the characteristic values of the additive observable A are also of order N. The exponential function therefore rapidly becomes negligible as soon as the value of A differs from the equilibrium value.

But if almost all the accessible volume corresponds to the equilibrium value of the observable A, we can calculate this value simply by taking the average of $A(x)$ over the entire accessible region of the phase space:

$$A_{\mathrm{eq}} = \left\langle A(x)\right\rangle = \frac{1}{|\Gamma|} \int_{\Gamma} \mathrm{d}x\, \delta\left(A(x) - A\right). \tag{3.64}$$

This gives us a very convenient "recipe" for calculating thermodynamic quantities. Let us observe that the average appearing in eq. (3.64) corresponds to the average over a probability distribution that is uniform within Γ and vanishes outside—this distribution is called **microcanonical**. In statistical mechanics, one often uses the word ensemble as a synonym for probability distribution in the phase space. The ensemble we have described, in which the internal energy has a fixed value, is called the **microcanonical ensemble** and was introduced by Gibbs.

Let us remark that conceptually the ensemble is nothing more than a rule to calculate the thermodynamic value of the observables, but that any other distribution, as long as it is fairly smooth within the phase space and only parametrized by the observables, would lead to equivalent results. We will see some additional examples further on.

3.10 The Maxwell Distribution

Let us consider for example the distribution of velocity in the ideal gas. We define the observable *number of particles having a certain value of velocity*:

$$N_v(\boldsymbol{v}) = \sum_{i=1}^{N} \delta\left(\frac{\boldsymbol{p}_i}{m} - \boldsymbol{v}\right). \tag{3.65}$$

Let us calculate the average value of $N_v(\boldsymbol{v})$ over the accessible region of phase space. Then $\langle N_v(\boldsymbol{v})\rangle \, \mathrm{d}\boldsymbol{v}$ is the average number of particles whose velocities belong to a small region of sides $\mathrm{d}\boldsymbol{v}$ around $\boldsymbol{v}$. Since the contribution of each particle to the average is the same, we obtain

$$\langle N_v(\boldsymbol{v})\rangle = N\left\langle \delta\left(\frac{\boldsymbol{p}_1}{m} - \boldsymbol{v}\right)\right\rangle. \tag{3.66}$$

We thus obtain, integrating over the positions $(\boldsymbol{r}_i)$ of the particles,

$$\langle N_v(\boldsymbol{v})\rangle = \frac{NV^N}{|\Gamma|}\int \mathrm{d}\boldsymbol{p}^N \, \delta\left(\sum_j \frac{p_j^2}{2m} - E\right) \delta\left(\frac{\boldsymbol{p}_1}{m} - \boldsymbol{v}\right). \tag{3.67}$$

The integral represents the projection of the $3N$-dimensional sphere defined by $\sum_i p_i^2 = 2mE$ on the three axes $p_{\alpha,1}$, $\alpha \in \{x, y, z\}$. Such projections are discussed in appendix E.3. We have in fact

$$\begin{aligned}\int \prod_{i=1}^{N} \mathrm{d}\boldsymbol{p}_i \, \delta\left(\sum_{j=1}^{N} \frac{p_j^2}{2m} - E\right) \delta\left(\frac{\boldsymbol{p}_1}{m} - \boldsymbol{v}\right) \\ = \int \prod_{i=2}^{N} \mathrm{d}\boldsymbol{p}_i \, \delta\left(\sum_{j=2}^{N} \frac{p_j^2}{2m} - E - \frac{1}{2}mv^2\right).\end{aligned} \tag{3.68}$$

This is the surface of a sphere in $3(N-1)$ dimensions of radius $(2mE - m^2v^2)^{1/2}$. Using the methods of appendix E.3 we obtain

$$\langle N_v(\boldsymbol{v})\rangle = \text{const.} \, \exp\left(-\frac{mv^2}{2E/3N}\right), \tag{3.69}$$

where we have neglected the difference between $3(N-1)$ and $3N$. The proportionality constant can be evaluated by the condition $\int \mathrm{d}\boldsymbol{v} \, \langle N_v(\boldsymbol{v})\rangle = N$. We thus obtain

$$\langle N_v(\boldsymbol{v})\rangle = N\left(\frac{3Nm}{2\pi E}\right)^{3/2} \exp\left(-\frac{mv^2}{2E/3N}\right), \tag{3.70}$$

which corresponds to the Maxwell distribution (1.9), if we take into account eq. (3.58) and the relation

$$\phi(v) = \frac{1}{N}\int \mathrm{d}\boldsymbol{v} \, \delta(v - |\boldsymbol{v}|) \, \langle N_v(\boldsymbol{v})\rangle = \frac{1}{N} 4\pi v^2 \, \langle N_v(v)\rangle. \tag{3.71}$$

3.11 The Ising Paramagnet

Let us now consider a simple model of paramagnet. We can disregard the molecules' translational degrees of freedom as irrelevant to our purposes, and assume that the molecules are arranged on a lattice. Let us also assume that each molecule possesses an elementary magnetic moment of magnitude μ_{m} that can be arranged (for quantum reasons) parallel or antiparallel to the direction of the z axis. Let us additionally assume that we can disregard all interactions between these magnetic moments and, as a matter of convenience, that we have rescaled the magnetic field h by the magnetic permeability of the vacuum μ_0. These assumptions result in the following definition of the model called the Ising paramagnet:

- Each molecule at site i of the lattice is associated with a variable σ_i, which can assume the values $+1$ and -1, depending on whether the corresponding magnetic moment is parallel or antiparallel to the z axis.
- The energy is only due to the interaction with the magnetic field h (which we assume is parallel to the positive z axis) and is equal to $-h\sum_{i=1}^{N}\mu_{\mathrm{m}}\sigma_i$.

Therefore, given the value E of the energy, the number L of molecules whose spin is oriented parallel to the z axis is given by

$$L=\frac{1}{2}\left(N-\frac{E}{\mu_{\mathrm{m}}h}\right). \tag{3.72}$$

The number $\mathcal{N}$ of ways in which these L molecules can be chosen is given by

$$\mathcal{N}=\frac{N!}{L!(N-L)!}. \tag{3.73}$$

By using Stirling's formula and introducing $m=\langle\sigma\rangle=-E/(N\mu_{\mathrm{m}}h)$, we have

$$\begin{aligned} S&=k_{\mathrm{B}}\ln\mathcal{N}\\ &\simeq -Nk_{\mathrm{B}}\left\{\frac{1}{2}(1-m)\ln\left[\frac{1}{2}(1-m)\right]+\frac{1}{2}(1+m)\ln\left[\frac{1}{2}(1+m)\right]\right\}. \end{aligned} \tag{3.74}$$

By taking the derivative of S with respect to E, we obtain

$$\frac{1}{T}=\frac{\partial m}{\partial E}\frac{\partial S}{\partial m}=\frac{k_{\mathrm{B}}}{\mu_{\mathrm{m}}}\cdot\frac{1}{2}\ln\frac{1+m}{1-m}=\frac{k_{\mathrm{B}}}{\mu_{\mathrm{m}}}\tanh^{-1}m, \tag{3.75}$$

which yields the equation of state for the magnetization M:

$$M=N\mu_{\mathrm{m}}m=N\mu_{\mathrm{m}}\tanh\left(\frac{\mu_{\mathrm{m}}h}{k_{\mathrm{B}}T}\right). \tag{3.76}$$

We have, of course, $E=-hM$.

Figure 3.3 shows the equation of state (3.76) m as a function of $\mu_{\mathrm{m}}h/k_{\mathrm{B}}T$. One sees that for $|\mu_{\mathrm{m}}h/k_{\mathrm{B}}T|\ll 1$ one has $m\propto h$, while for $|\mu_{\mathrm{m}}h/k_{\mathrm{B}}T|\gg 1$, m tends to saturate at the values $m=\pm 1$.

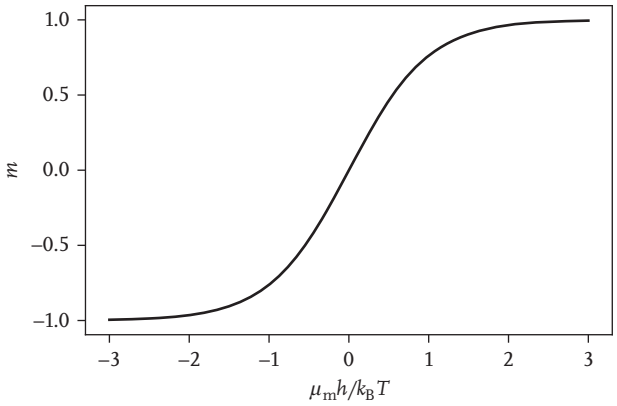


FIGURE 3.3. Equation of state of the Ising paramagnet: $m = M/N\mu_m$ as a function of $\mu_m h/k_B T$.

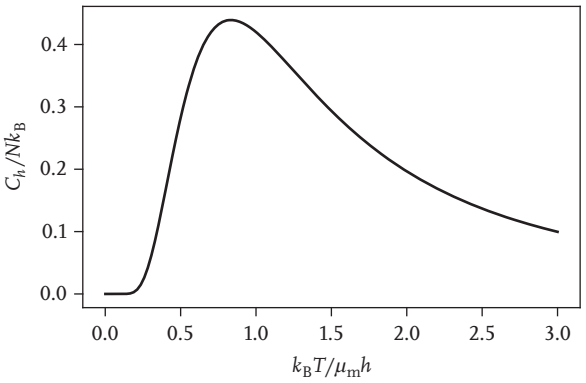


FIGURE 3.4. Specific heat of the Ising paramagnet. C_h/Nk_B is plotted versus $k_B T/\mu_m h$.

The specific heat with magnetic field held constant can be obtained by taking the derivative of the energy with respect to T:

$$C_h = \left.\frac{\partial E}{\partial T}\right)_h = -N\mu_m h \frac{\partial m}{\partial T} = Nk_B \left(\frac{\mu_m h}{k_B T}\right)^2 \left[1 - \tanh^2\left(\frac{\mu_m h}{k_B T}\right)\right]. \tag{3.77}$$

Figure 3.4 shows the behavior of C_h/Nk_B as a function of $k_B T/\mu_m h$. One notices that the specific heat vanishes exponentially for $T \to 0$, whereas it decreases as T^{-2} for large values of T. It exhibits a maximum for $k_B T \simeq 0.8336\ |\mu_m h|$.

Let us remark that eq. (3.75) assigns negative values to T for $m < 0$, that is, when the majority of the spins are antiparallel to the magnetic field. Indeed, if we plot the entropy per spin S/Nk_B as a function of m (fig. 3.5), we see that it increases from 0 to $\ln 2$ for $-1 \le m < 0$, and decreases from $\ln 2$ to 0 for $0 < m \le 1$. Since $E = -N\mu_m h$, the entropy decreases as E increases when $E > 0$, that is, $m < 0$, in other words, when the energy is larger than its value for $T \to \infty$. This implies that negative temperatures are *hotter* than infinite temperatures! We can expect this situation to arise in all systems in which the range of possible energy values is bounded above. In this case, there is necessarily an

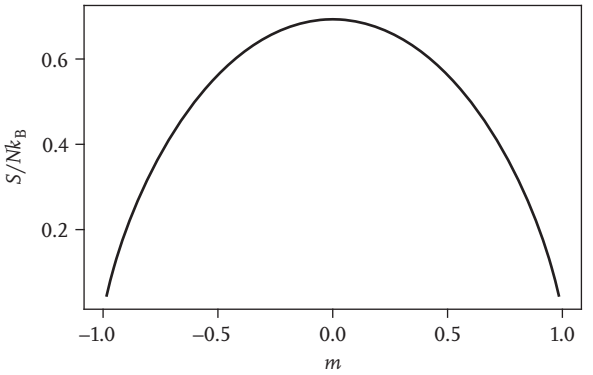


FIGURE 3.5. Entropy per spin S/Nk_B of the Ising paramagnet as a function of $m = M/N\mu_m$.

entropy maximum in the interior of this range. It is clear that in this situation the expression (3.15) is no more equivalent to the expression (3.14) that we have used above. Indeed, $S(E)$, given by eq. (3.15) remains fixed at the maximum value when E increases beyond the location E^* of the maximum.

Are these situations physically realizable? Particle systems possess a kinetic energy, whose values are not bounded above. Therefore, negative-temperature states cannot appear in them. On the other hand, a system of "pure spins" could exhibit a negative-temperature state if it were possible to prepare it in a state with energy $E > E^*$. This requires isolating the spin degrees of freedom from the translational ones. In some systems the coupling of the nuclear spins with the lattice is very weak at low temperatures, and it is possible to consider the spin system as almost adiabatically isolated from the translational degrees of freedom of the lattice. If we then first equilibrate the system with a certain magnetic field h, and then we suddenly reverse h, we can obtain a negative-temperature state. Of course, such a state has a finite lifetime, since the spin system will leak heat to the lattice (after all, its temperature is higher than that of the lattice) and its energy will therefore slowly decrease. An experimental realization of this system is described in [Purc51]. A thorough discussion of the implications of negative temperatures is given in [Rams56], and a recent comprehensive review is given in [Bald21].

Exercise 3.2 (The Bohr-van Leeuwen theorem) Prove that when statistical mechanics and classical mechanics are applied consistently, and the particles have no intrinsic magnetic moment, the thermal average of the magnetization is always zero.

1. Express the Hamiltonian $\mathcal{H}(\boldsymbol{p}, \boldsymbol{r})$ of a system of N charged particles as a function of the coordinates $\boldsymbol{r} = (\boldsymbol{r}_i)$ and the momenta $\boldsymbol{p} = (\boldsymbol{p}_i)$ of the particles, and of the vector potential $\boldsymbol{A}(\boldsymbol{r})$ of the field, such that the magnetic induction field $\boldsymbol{B} = \nabla \times \boldsymbol{A}$.

2. Evaluate the magnetization $\boldsymbol{M} = -\langle \partial \mathcal{H} / \partial \boldsymbol{B} \rangle$ in the canonical ensemble.

Thus diamagnetism in materials have a quantum origin. This theorem was discovered by N. Bohr in his doctoral dissertation, and rediscovered a few years later by H. J. van Leeuwen.

3.12 The Canonical Ensemble

The microcanonical ensemble that we have described in the preceding sections corresponds to the most direct description of thermodynamic systems, in which the thermodynamic equilibrium states are identified by the values of a collection of extensive thermodynamic variables. We have seen that in thermodynamics it is possible to provide a different but equivalent description of the same thermodynamic state by considering as fixed the values of some intensive variables, rather than those of their conjugate extensive ones. In an analogous way, for each such representation of thermodynamic equilibrium states, it is possible to give an appropriate and equivalent description in statistical mechanics.

The most important case is that in which the intensive variable we want to introduce is the temperature and the corresponding intensive one, which we wish to remove, is the internal energy. In this, case, we suppose that we put the system we are interested in, which we denote by S and call "system," in contact with a much larger system, which we denote by R and which we call "reservoir." System and reservoir can freely exchange energy. At equilibrium, they will have the same temperature—since the reservoir is much larger than the system, however, we can assume that, whatever the amount of energy it exchanges with the system, its temperature does not vary. In this fashion, we have imposed a fixed temperature value on the system. We wish to calculate the average A_{eq} of any observable A relative to the system S only. In order to do this, we observe that the state of the global system $(\mathrm{S} \cup \mathrm{R})$ is defined by the pair $(x_{\mathrm{S}}, x_{\mathrm{R}})$, where x_{S} is the representative point of the system proper, and x_{R} that of the reservoir. Therefore, by applying the rule we derived in the preceding chapter, we obtain:

$$A_{\mathrm{eq}} = \langle A \rangle = \frac{1}{|\Gamma|} \int_{\Gamma} \mathrm{d}x_{\mathrm{S}}\, \mathrm{d}x_{\mathrm{R}}\, A(x_{\mathrm{S}}). \tag{3.78}$$

The region Γ is defined by the value E of the total internal energy and by the values of the other extensive variables that we will not specify. Let us rewrite the integral as follows:

$$\langle A \rangle = \frac{1}{|\Gamma|} \int \mathrm{d}x_{\mathrm{S}}\, \mathrm{d}x_{\mathrm{R}}\, A(x_{\mathrm{S}})\, \delta\left(\mathcal{H}^{(\mathrm{S})}(x_{\mathrm{S}}) + \mathcal{H}^{(\mathrm{R})}(x_{\mathrm{R}}) - E\right). \tag{3.79}$$

This integral can be rearranged in the following manner:

$$\langle A \rangle = \frac{1}{|\Gamma|} \int \mathrm{d}x_{\mathrm{S}}\, A(x_{\mathrm{S}}) \times \int \mathrm{d}x_{\mathrm{R}}\, \delta\left(\mathcal{H}^{(\mathrm{R})}(x_{\mathrm{R}}) - \left(E - \mathcal{H}^{(\mathrm{S})}(x_{\mathrm{S}})\right)\right). \tag{3.80}$$

The integral in $\mathrm{d}x_{\mathrm{R}}$ that appears in this equation is the volume of phase space accessible to the reservoir when its internal energy assumes the value $E - \mathcal{H}^{(\mathrm{S})}(x_{\mathrm{S}})$. This quantity is expressed by the entropy:

$$\int \mathrm{d}x_{\mathrm{R}}\, \delta\left[\mathcal{H}^{(\mathrm{R})}(x_{\mathrm{R}}) - \left(E - \mathcal{H}^{(\mathrm{S})}(x_{\mathrm{S}})\right)\right] \simeq \exp\left[\frac{1}{k_{\mathrm{B}}} S^{(\mathrm{R})}\left(E - \mathcal{H}^{(\mathrm{S})}(x_{\mathrm{S}})\right)\right]. \tag{3.81}$$

Taking into account that the system is much smaller than the reservoir, we can expand the exponent to first order in a Taylor series:

$$\exp\left[\frac{1}{k_B}S^{(R)}\left(E-\mathcal{H}^{(S)}(x_S)\right)\right] \simeq \exp\left[\frac{1}{k_B}S^{(R)}(E)\right]\exp\left[-\frac{1}{k_B}\left.\frac{\partial S^{(R)}}{\partial E}\right|_E \mathcal{H}^{(S)}(x_S)\right]. \tag{3.82}$$

At this point, we note that $\partial S^{(R)}/\partial E = 1/T$, and that in our hypotheses, this quantity is fixed. We can then write:

$$\langle A\rangle = \frac{1}{Z}\int \mathrm{d}x_S\, A(x_S)\, \exp\left[-\frac{\mathcal{H}^{(S)}(x_S)}{k_B T}\right]. \tag{3.83}$$

The normalization constant Z is called the **partition function** and is given by

$$Z = \int \mathrm{d}x_S\, \exp\left[-\frac{\mathcal{H}^{(S)}(x_S)}{k_B T}\right]. \tag{3.84}$$

We see that we can now freely integrate over the system's entire phase space but that the different regions of the phase space are weighed proportionally to the factor

$$\exp\left[-\frac{\mathcal{H}^{(S)}(x_S)}{k_B T}\right], \tag{3.85}$$

which is called the **Boltzmann factor**. We can interpret this factor as the probability density of a certain point in the phase space. In this way, we have defined a new probability distribution—in other words, a new ensemble, which is called the **canonical ensemble**. In order for these results to be compatible with those we obtained from the microcanonical ensemble, it is necessary that the contribution to $\langle A\rangle$ be dominated by the region of phase space in which the system's internal energy assumes the equilibrium value. In order to prove that this is the case, let us rearrange the integral according to the values of internal energy. From here on, we drop the S indices, because we are examining explicitly only the system itself. We therefore obtain

$$\langle A\rangle = \frac{1}{Z}\int \mathrm{d}E\, \exp\left(-\frac{E}{k_B T}\right)\int \mathrm{d}x\, A(x)\,\delta\left(\mathcal{H}(x)-E\right). \tag{3.86}$$

The integral in x is proportional to the average microcanonical value of A in the state in which the internal energy is equal to E, and we denote it with $A_{eq}(E)$. The proportionality factor is the corresponding volume $\left|\Gamma(E)\right|$ of phase space, which we can express by means of the entropy: $\left|\Gamma(E)\right| \simeq \mathrm{e}^{S(E)/k_B}$. The expression of $\langle A\rangle$ therefore becomes

$$\langle A\rangle = \frac{1}{Z}\int \mathrm{d}E\, A_{eq}(E)\, \exp\left[-\frac{E-TS(E)}{k_B T}\right]. \tag{3.87}$$

Since $A_{eq}(E)$ varies slowly, the integral is dominated by the behavior of the exponential, in which the extensive functions E and S appear. If we evaluate it by means of the saddle-point

method (see appendix E), we notice that it is dominated by a region in which $E \simeq E_{\mathrm{eq}}$, where E_{eq} is the extremal of $E - TS(E)$ and therefore satisfies

$$\left.\frac{\partial S}{\partial E}\right|_{E_{\mathrm{eq}}} = \frac{1}{T}. \tag{3.88}$$

We can thus identify E_{eq} with the thermodynamic equilibrium value of the energy—in other words, with the internal energy. The contribution of the other values of E decreases rapidly as soon as $\left|E - E_{\mathrm{eq}}\right|$ becomes of the order of $N^{1/2}$ or larger. Considering the integral that defines the partition function, we have

$$Z \simeq \exp\left[-\frac{E_{\mathrm{eq}} - TS(E_{\mathrm{eq}})}{k_{\mathrm{B}}T}\right] = \exp\left(-\frac{F}{k_{\mathrm{B}}T}\right), \tag{3.89}$$

which corresponds to

$$F = -k_{\mathrm{B}}T \ln Z, \tag{3.90}$$

where F is the Helmholtz free energy expressed in the relevant natural variables, that is, the temperature T and the other extensive variables X. Knowledge of the partition function therefore allows us to obtain the thermodynamic properties of the system.

3.12.1 *Energy Fluctuations*

More particularly, the value E of the internal energy can be obtained by differentiating $\ln Z(\beta) = \ln \int \mathrm{d}x\, \mathrm{e}^{-\beta \mathcal{H}(x)}$ with respect to β and then setting $\beta = 1/k_{\mathrm{B}}T$:

$$\left.\frac{\partial \ln Z(\beta)}{\partial \beta}\right|_{\beta = 1/k_{\mathrm{B}}T} = -\frac{1}{Z}\int \mathrm{d}x\, \mathcal{H}(x)\, \exp\left[-\frac{\mathcal{H}(x)}{k_{\mathrm{B}}T}\right] = -\langle \mathcal{H} \rangle = -E. \tag{3.91}$$

Taking the second derivative of $\ln Z$ we obtain

$$\begin{aligned}\left.\frac{\partial^2 \ln Z(\beta)}{\partial \beta^2}\right|_{\beta = 1/k_{\mathrm{B}}T} &= \frac{1}{Z}\int \mathrm{d}x\, \mathcal{H}^2(x)\, \exp\left[-\frac{\mathcal{H}(x)}{k_{\mathrm{B}}T}\right] \\ &\quad - \frac{1}{Z^2}\left\{\int \mathrm{d}x\, \mathcal{H}(x)\, \exp\left[-\frac{\mathcal{H}(x)}{k_{\mathrm{B}}T}\right]\right\}^2 \\ &= \left\langle \mathcal{H}^2 \right\rangle - \langle \mathcal{H} \rangle^2 .\end{aligned} \tag{3.92}$$

In this manner, we have obtained the variance of internal energy in the canonical ensemble. It is proportional to N. Moreover, since it cannot be negative, it follows that $\ln Z$ must be a convex function of its argument $1/k_{\mathrm{B}}T$.

It is easy to relate this result to the stability properties of Helmholtz free energy. In fact, by taking the derivative of eq. (3.91), we obtain

$$\mathrm{Var}\, \mathcal{H} = \left\langle \mathcal{H}^2 \right\rangle - \langle \mathcal{H} \rangle^2 = -\frac{\partial E}{\partial (1/k_{\mathrm{B}}T)} = k_{\mathrm{B}}T^2 \frac{\partial E}{\partial T} = k_{\mathrm{B}}T^2 C, \tag{3.93}$$

where C is the specific heat. We have thus related a statistical quantity (variance), connected with the fluctuations of internal energy, to a thermodynamic quantity (specific heat). Since the variance is positive, also the specific heat is positive. This is the first of a series of important relations that come under the general designation of **fluctuation-dissipation theorems**.

3.12.2 The Ideal Gas in the Canonical Ensemble

The advantage of using the canonical ensemble becomes obvious when we are dealing with systems whose Hamiltonian has the form

$$\mathcal{H}(x) = \sum_i h(x_i). \tag{3.94}$$

In other words, it decomposes into terms, each of which is a function of only one or a few degrees of freedom. In fact, while in the microcanonical ensemble, we must satisfy a *global* constraint in which all the degrees of freedom appear; in the canonical ensemble, we can split our integrals into factors in which only a few degrees of freedom appear at each time, and which can be evaluated explicitly.

The ideal monoatomic gas obviously falls within this category. We have in fact

$$\mathcal{H} = \sum_{i=1}^{N} \left(\frac{p_i^2}{2m} + v(\boldsymbol{r}_i) \right) = \sum_i \left(\sum_\alpha \frac{p_{\alpha,i}^2}{2m} + v(\boldsymbol{r}_i) \right), \tag{3.95}$$

where $v(\boldsymbol{r})$ is the confining potential (3.53). Let us then calculate the partition function:

$$Z = \frac{1}{N!} \int \prod_\alpha \prod_{i=1}^{N} \left(\frac{\mathrm{d}p_{\alpha,i}\, \mathrm{d}r_{\alpha,i}}{h} \right) \exp \left\{ -\frac{1}{k_\mathrm{B} T} \left[\sum_i \left(\sum_\alpha \frac{p_{\alpha,i}^2}{2m} + v(\boldsymbol{r}_i) \right) \right] \right\}. \tag{3.96}$$

The factor $1/N!$ comes from the indistinguishability of the particles, as in the microcanonical case. The integral over the positions yields V^N as in the microcanonical case, while the integral over momenta factorizes into $3N$ identical integrals, yielding

$$Z = \frac{V^N}{N!} \left[\int \frac{\mathrm{d}p}{h} \exp \left(-\frac{p^2}{2mk_\mathrm{B} T} \right) \right]^{3N} = \frac{V^N}{N!} \left(\frac{2\pi mk_\mathrm{B} T}{h^2} \right)^{3N/2} = \frac{V^N}{N! \lambda_\mathrm{B}^{3N}}, \tag{3.97}$$

where we have introduced the **thermal de Broglie wavelength**

$$\lambda_\mathrm{B} = \left(\frac{h^2}{2\pi mk_\mathrm{B} T} \right)^{1/2}. \tag{3.98}$$

By taking the logarithm and using Stirling's formula, we obtain

$$\ln Z \simeq N \ln \left[\frac{\mathrm{e} V}{N} \left(\frac{2\pi mk_\mathrm{B} T}{h^2} \right)^{3/2} \right] = N \ln \left(\frac{\mathrm{e}}{\rho \lambda_\mathrm{B}^3} \right). \tag{3.99}$$

In this equation we have introduced the numerical density $\rho = N/V$. From this equation, we obtain the Helmholtz free energy F and, by taking derivatives, the equations

of state:

$$F = -k_B T \ln Z = -N k_B T \ln\left(\frac{e}{\rho \lambda_B^3}\right); \tag{3.100}$$

$$S = -\left.\frac{\partial F}{\partial T}\right)_V = N k_B \left[\ln\left(\frac{e}{\rho \lambda_B^3}\right) - \frac{3}{2}\right]; \tag{3.101}$$

$$p = -\left.\frac{\partial F}{\partial V}\right)_T = k_B T \rho = \frac{N k_B T}{V}. \tag{3.102}$$

From the definition $F = E - TS$ or, equivalently, by taking the derivative of $\ln Z$ with respect to $\beta = 1/k_B T$, we obtain the expression of the internal energy:

$$E = \frac{3}{2} N k_B T. \tag{3.103}$$

By taking a further derivative of this result, we obtain the specific heat at constant volume:

$$C_V = \left.\frac{\partial E}{\partial T}\right)_V = \frac{3}{2} N k_B. \tag{3.104}$$

As a consequence, by Mayer's relation (2.158), the ratio $\gamma = C_p/C_V$ of the specific heat at constant pressure vs. that at constant volume is equal to 5/3. Then, by eq. (2.160), the product $pV^{5/3}$ remains constant for a monoatomic gas in an adiabatic transformation.

3.13 Generalized Ensembles

Instead of allowing the value of the internal energy E to fluctuate, keeping the value of the temperature T fixed, we can allow fluctuations in a generic extensive variable X (other than the energy). We can then follow the procedure of the preceding section. We imagine that we put our system S (the system proper) in contact with a much larger system R (the reservoir). The reservoir can exchange the quantity X with the system and is characterized by the value f of the corresponding conjugate variable. (We make the unphysical assumption that this exchange does not involve the energy.) Keeping in mind that f is determined by the relation

$$\left.\frac{\partial S^{(R)}}{\partial X^{(R)}}\right)_E = -\frac{f}{T}, \tag{3.105}$$

we obtain

$$S^{(R)}(X - X^{(S)}) = S^{(R)}(X) + \frac{f X^{(S)}}{T}. \tag{3.106}$$

We obtain therefore the following expression of $\langle A \rangle$:

$$\langle A \rangle = \frac{1}{|\Gamma|} \int_\Gamma dx_S \, dx_R \, A(x_S)$$

$$
\begin{aligned}
&= \frac{1}{|\Gamma|} \int \mathrm{d}x_\mathrm{S}\, A(x_\mathrm{S}) \int \mathrm{d}x_\mathrm{R}\, \delta\left(X^{(\mathrm{S})}(x_\mathrm{S}) + X^{(\mathrm{R})}(x_\mathrm{R}) - X\right) \\
&= \frac{1}{|\Gamma|} \int \mathrm{d}x_\mathrm{S}\, A(x_\mathrm{S})\, \exp\left[\frac{1}{k_\mathrm{B}} S^{(\mathrm{R})}(X - X^{(\mathrm{S})}(x_\mathrm{S}))\right] \\
&\simeq \frac{1}{Z} \int \mathrm{d}x_\mathrm{S}\, A(x_\mathrm{S})\, \exp\left[\frac{f X^{(\mathrm{S})}(x_\mathrm{S})}{k_\mathrm{B} T}\right],
\end{aligned}
\tag{3.107}
$$

where Z is a normalization constant. We can now see that the different points in phase space are weighted by a probability distribution

$$
P(x) = \frac{1}{Z} \exp\left[\frac{f X^{(\mathrm{S})}(x_\mathrm{S})}{k_\mathrm{B} T}\right]. \tag{3.108}
$$

We have thus introduced a new ensemble, parametrized by the intensive variable f.

It is a consequence of the principles of thermodynamics that practically the entire weight is concentrated in the region of phase space in which X assumes the equilibrium value. One can become convinced of this fact by again following the line of reasoning of sec. 3.12.

We now focus on the system proper and drop the index S. The normalization constant Z can be easily interpreted in thermodynamic terms. We have

$$
Z = \int \mathrm{d}x\, \exp\left[\frac{f X(x)}{k_\mathrm{B} T}\right] = \int \mathrm{d}X\, \exp\left(\frac{f X}{k_\mathrm{B} T}\right) \int \mathrm{d}x\, \delta(X - X(x)). \tag{3.109}
$$

The integral over phase space in this expression is equal to the entropy expressed as a function of the extensive variable X (as well as of the other extensive thermodynamic variables). Therefore,

$$
Z = \int \mathrm{d}X\, \exp\left[\frac{f X + T S(X)}{k_\mathrm{B} T}\right]. \tag{3.110}
$$

Since both S and X are extensive variables (and are therefore proportional to N), we can calculate this integral with the saddle point method, and we obtain

$$
Z \simeq \exp\left[\frac{f X_\mathrm{eq} + T S(X_\mathrm{eq})}{k_\mathrm{B} T}\right], \tag{3.111}
$$

where X_eq is the value of X for which the exponent becomes a maximum—in other words, for which

$$
\left.\frac{\partial S}{\partial X}\right|_{X_\mathrm{eq}} = -\frac{f}{T}. \tag{3.112}
$$

By the same token, $X_\mathrm{eq} = \langle X \rangle$, where $\langle X \rangle$ is the average of X over the probability distribution (3.108). We therefore see that the argument of the exponential is a Legendre transform of the entropy. These thermodynamic functions are known as **Massieu functions**—we

will see, however, that in the most useful cases we will end up once again with the usual Legendre tranforms of the internal energy that we saw in chapter 2.

3.13.1 Average Values and Fluctuations

The normalization constant allows us to obtain the values for the quantities of interest by simple differentiation. One therefore has

$$\langle X\rangle = \frac{1}{Z}\int \mathrm{d}x\, X(x)\, \exp\left[\frac{fX(x)}{k_\mathrm{B}T}\right] = \frac{\partial \ln Z}{\partial(f/k_\mathrm{B}T)}. \tag{3.113}$$

By taking the derivative of this expression, we obtain

$$\frac{\partial^2 \ln Z}{\partial(f/k_\mathrm{B}T)^2} = \left\langle X^2\right\rangle - \langle X\rangle^2 = \mathrm{Var}\, X = k_\mathrm{B}T^2\frac{\partial\langle X\rangle}{\partial f}, \tag{3.114}$$

which is another fluctuation-dissipation relation. Since the variance of X is positive, the derivative $\partial\langle X\rangle/\partial f$ must also be positive. This is also a consequence of stability, that is, of the concavity of the entropy.

In this relation, Boltzmann's constant appears as a proportionality constant between the variance of X and the thermodynamic derivative $\partial\langle X\rangle/\partial T$—therefore it is in principle possible to measure Boltzmann's constant (and thus Avogadro's number) by measuring the fluctuations of the thermodynamic variables. However, for systems of macroscopic sizes, these fluctuations are well-nigh unobservable. To proceed in this direction, one has to turn to small systems, or to situations where the fluctuations are much enhanced.

3.13.2 The Paramagnet

In the case of the paramagnet, we want to switch from the ensemble with a fixed magnetization M to one in which the magnetic field h is fixed. Denote by $\sigma = (\sigma_i)$ the generic configuration of the N spins. We then have

$$\begin{aligned} Z &= \sum_\sigma \exp\left[\frac{hM(\sigma)}{k_\mathrm{B}T}\right] = \sum_\sigma \exp\left(\frac{1}{k_\mathrm{B}T}\sum_{i=1}^N \mu_\mathrm{m} h\sigma_i\right) \\ &= \prod_{i=1}^N\left[\sum_{\sigma_i}\exp\left(\frac{\mu_\mathrm{m}h\sigma_i}{k_\mathrm{B}T}\right)\right] = \left(\mathrm{e}^{\mu_\mathrm{m}h/k_\mathrm{B}T} + \mathrm{e}^{-\mu_\mathrm{m}h/k_\mathrm{B}T}\right)^N \\ &= \left[2\cosh\left(\frac{\mu_\mathrm{m}h}{k_\mathrm{B}T}\right)\right]^N. \end{aligned} \tag{3.115}$$

The corresponding free energy is given by

$$F = -k_\mathrm{B}T\ln Z = -Nk_\mathrm{B}T\ln\left[2\cosh\left(\frac{\mu_\mathrm{m}h}{k_\mathrm{B}T}\right)\right]. \tag{3.116}$$

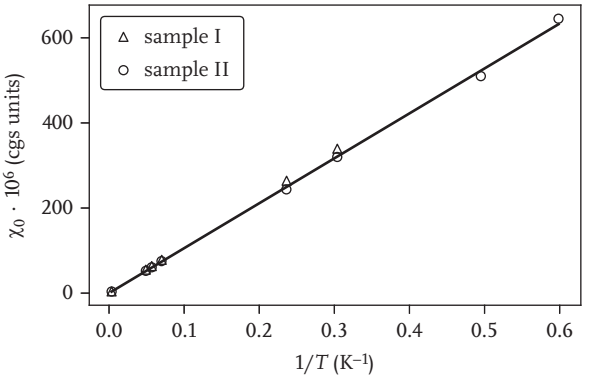


FIGURE 3.6. Illustration of Curie's law. Data on copper potassium sulphate hexahydrate [Hups42]. The line is a fit of $\chi = \kappa/T$ to the data.

The average magnetization is given by

$$\langle M \rangle = \frac{\partial \ln Z}{\partial (h/k_{\mathrm{B}}T)} = N\mu_{\mathrm{m}} \tanh\left(\frac{\mu_{\mathrm{m}} h}{k_{\mathrm{B}}T}\right), \tag{3.117}$$

in agreement with what has been obtained from the ensemble with fixed M. The variance of M is given by

$$\langle M \rangle^2 - \langle M \rangle^2 = \frac{\partial^2 \ln Z}{\partial (h/k_{\mathrm{B}}T)^2} = N\mu_{\mathrm{m}}^2 \left[1 - \tanh^2\left(\frac{\mu_{\mathrm{m}} h}{k_{\mathrm{B}}T}\right)\right] = k_{\mathrm{B}}T\,\chi, \tag{3.118}$$

where χ is the isothermal susceptibility:

$$\chi = \left.\frac{\partial \langle M \rangle}{\partial h}\right)_T. \tag{3.119}$$

We have thus obtained another fluctuation-dissipation relation. From this relation, we we obtain $\partial M/\partial h)_T = N\mu_{\mathrm{m}}^2/k_{\mathrm{B}}T$ for $h = 0$, that is, that $\chi(h{=}0)$ is inversely proportional to the absolute temperature—this result is known as **Curie's law** [Curi95] and is the basis of some methods used to measure low temperatures. An illustration is given in fig. 3.6.

Exercise 3.3 (Langevin theory of paramagnetism) Suppose that the system is made by N *classical* magnetic dipoles of magnitude μ, whose direction is identified by the spherical angles θ (latitude) and ϕ (longitude). The dipoles have a moment of inertia I for rotations normal to their axis.

1. Write the classical Lagrangian for such a system, supposing that the potential energy derives only from the interaction with an external magnetic field of magnitude B directed along the z axis. Obtain the corresponding Hamiltonian.

2. Evaluate the canonical partition function of the system by a suitable phase-space integral.

3. Obtain the isotherm of the system, that is, the average magnetization per spin $m(B)$ at a given temperature T. Check if the Curie law applies to it.

Exercise 3.4 (An elementary model of elasticity) Consider a fiber made by N molecules of length ℓ_0 and width $\ell_1 > \ell_0$, bound one after the other along the x axis. Each molecule can be in two positions: either normal to the x axis or parallel to it. In the first case, it occupies a length equal to ℓ_0, and in the second, a length equal to ℓ_1. The first molecule is fixed at the origin, and on the last one is applied a force F that pulls in the direction of the positive x axis. We denote by L the abscissa of the last molecule.

1. Write the Hamiltonian of the system, as a function of the configuration $\tau = (\tau_1, \dots, \tau_N)$, where $\tau_i = 0$ if the molecule is normal to the x axis, and $\tau_i = 1$ if it is parallel to it.
2. Evaluate the equation of state $L = L(N, F, T)$. Deduce that the system follows Hooke's law for small values of F, and evaluate the Hooke constant.

3.14 The p-T Ensemble

Let us now suppose that we fix the values of the temperature T and of the pressure p and let therefore the internal energy and the volume fluctuate. The corresponding ensemble is called the p-T ensemble and is described by the following expression:

$$\langle A \rangle = \frac{1}{Z} \int \mathrm{d}x\, A(x)\, \exp\left[-\frac{\mathcal{H}(x) + pV(x)}{k_\mathrm{B} T} \right]. \tag{3.120}$$

The partition function is connected to the Gibbs free energy:

$$\ln Z = -\frac{E - TS + pV}{k_\mathrm{B} T} = -\frac{G}{k_\mathrm{B} T}. \tag{3.121}$$

We can easily obtain the following relations for the average values of the internal energy and the volume:

$$E = \langle \mathcal{H} \rangle = -\left. \frac{\partial \ln Z}{\partial (1/k_\mathrm{B} T)} \right)_{p/k_\mathrm{B} T}; \tag{3.122}$$

$$\langle V \rangle = -\left. \frac{\partial \ln Z}{\partial (p/k_\mathrm{B} T)} \right)_{1/k_\mathrm{B} T}. \tag{3.123}$$

By taking further derivatives, we obtain

$$\begin{aligned} \mathrm{Var}\, \mathcal{H} = \left\langle \mathcal{H}^2 \right\rangle - \langle \mathcal{H} \rangle^2 &= -\left. \frac{\partial E}{\partial (1/k_\mathrm{B} T)} \right)_{p/k_\mathrm{B} T} \\ &= k_\mathrm{B} T^2 \left[\left. \frac{\partial E}{\partial T} \right)_p + \frac{p}{T} \left. \frac{\partial E}{\partial p} \right)_T \right]. \end{aligned} \tag{3.124}$$

We now show that the energy fluctuations are larger in the p-T ensemble than in the canonical one. This is due to the fact that some energy fluctuations are induced by volume

fluctuations. We have in fact

$$\operatorname{Var} V = \langle \Delta V^2 \rangle = \frac{\partial^2 \ln Z}{\partial (p/k_B T)^2} = -k_B T \left. \frac{\partial \langle V \rangle}{\partial p} \right)_T = k_B T \langle V \rangle \kappa_T, \tag{3.125}$$

where

$$\kappa_T = -\frac{1}{\langle V \rangle} \left. \frac{\partial \langle V \rangle}{\partial p} \right)_T \tag{3.126}$$

is the isothermal compressibility.

We obtain therefore, using eq. (3.93) and the result just obtained,

$$\begin{aligned} \operatorname{Var} \mathcal{H} &= \langle \Delta \mathcal{H}^2 \rangle_{p\text{-}T} = \langle \Delta \mathcal{H}^2 \rangle_C + \left. \frac{\partial E}{\partial V} \right)_T^2 \langle \Delta V^2 \rangle_{p\text{-}T} \\ &= k_B T^2 \left[\left. \frac{\partial E}{\partial T} \right)_V - \frac{1}{T} \left. \frac{\partial E}{\partial V} \right)_T^2 \left. \frac{\partial V}{\partial p} \right)_T \right] \geq k_B T^2 C_V, \end{aligned} \tag{3.127}$$

where $\langle \ldots \rangle_{p\text{-}T}$ refers to the p-T ensemble and $\langle \ldots \rangle_C$ to the canonical ensemble with the same value of $\langle V \rangle$.

Exercise 3.5 Making use of the Maxwell relations and the expression $p = -\partial F / \partial V)_T$ of the pressure, prove that the two expressions of $\operatorname{Var} \mathcal{H}$ in eqs. (3.124) and (3.127) are indeed equal.

Last, we have the mixed derivatives:

$$\begin{aligned} \operatorname{Corr}(\mathcal{H}, V) &= \langle \mathcal{H}\, V \rangle - \langle \mathcal{H} \rangle \langle V \rangle = \frac{\partial^2 \ln Z}{\partial (1/k_B T) \partial (p/k_B T)} \\ &= k_B T^2 \left. \frac{\partial V}{\partial T} \right)_p = k_B T \left. \frac{\partial E}{\partial p} \right)_T . \end{aligned} \tag{3.128}$$

We see that the Maxwell relations are connected to the symmetry of correlations. Moreover, the convexity of the partition function is related to the fact that the correlation matrix is positive definite.

3.14.1 One-Dimensional Hard Sphere Gas

As an application of the p-T ensemble, we can solve our first problem with interacting particles. Let us suppose that we have some particles arranged on a straight line. Each particle is a hard sphere of diameter ℓ. Two particles that do not touch do not interact.

The potential energy of the particles is therefore given by

$$U(r_1, \ldots, r_N) = \frac{1}{2} \sum_{i \neq j} u(r_i - r_j), \tag{3.129}$$

where

$$u(r) = \begin{cases} \infty, & \text{if } |r| \leq \ell; \\ 0, & \text{otherwise.} \end{cases} \tag{3.130}$$

Calculating the partition function becomes very simple in the p-T ensemble.

Since the particles cannot exchange positions, we can imagine numbering them consecutively starting with the leftmost one. (This gets rid of the $1/N!$ factor.) We additionally set an impenetrable wall fixed at the origin, and let the particles move along the positive x axis. Then the volume of the system is equal to the coordinate of the last particle. We thus obtain the following expression of the partition function:

$$\begin{aligned} Z = \int \prod_{i=1}^{N} \frac{\mathrm{d}p_i}{h} \int_{\ell}^{\infty} \mathrm{d}r_1 \int_{r_1+\ell}^{\infty} \mathrm{d}r_2 \cdots \int_{r_{N-1}+\ell}^{\infty} \mathrm{d}r_N \\ \times \exp\left(-\frac{1}{k_\mathrm{B}T} \sum_{i=1}^{N} \frac{p_i^2}{2m} - \frac{p\, r_N}{k_\mathrm{B}T} \right). \end{aligned} \tag{3.131}$$

The momentum integrals yield λ_B^{-N}, where λ_B is the thermal de Broglie wavelength. The integral over the positions $(r_1, \ldots, r_N)$ is called the configuration integral. To evaluate it, let us change the integration variables to

$$\xi_1 = r_1; \qquad \xi_i = r_i - r_{i-1}, \quad i > 1. \tag{3.132}$$

We thus obtain the following expression of the partition function:

$$Z = \lambda_\mathrm{B}^{-N} \int_{\ell}^{\infty} \mathrm{d}\xi_1 \int_{\ell}^{\infty} \mathrm{d}\xi_2 \cdots \int_{\ell}^{\infty} \mathrm{d}\xi_N \, \exp\left(-\frac{p}{k_\mathrm{B}T} \sum_{i=1}^{N} \xi \right). \tag{3.133}$$

The integral factorizes into N identical integrals, yielding

$$Z = \lambda_\mathrm{B}^{-N} \left[\int_{\ell}^{\infty} \mathrm{d}\xi \, \exp\left(-\frac{p\,\xi}{k_\mathrm{B}T} \right) \right]^N = \left[\frac{1}{\lambda_\mathrm{B}} \frac{k_\mathrm{B}T}{p} \exp\left(-\frac{p\,\ell}{k_\mathrm{B}T} \right) \right]^N. \tag{3.134}$$

By taking the derivative of $\ln Z$ with respect to $p/k_\mathrm{B}T$, we obtain the equation of state:

$$V = -\frac{\partial \ln Z}{\partial (p/k_\mathrm{B}T)} = N \left(\frac{k_\mathrm{B}T}{p} + \ell \right). \tag{3.135}$$

In other words,

$$p\,(V - N\ell) = N k_\mathrm{B} T. \tag{3.136}$$

The internal energy is obtained by taking the derivative of $\ln Z$ with respect to $1/k_\mathrm{B}T$ (at $p/k_\mathrm{B}T$ constant) and is equal to $Nk_\mathrm{B}T/2$.

Exercise 3.6 Extend the preceding result to the case in which the potential $u(r)$ between the particles contains, beyond the hard-core repulsion, an attractive interaction

with a range between ℓ and 2ℓ (such that only close nearest-neighbor particles can interact).

3.15 Quantum Ensembles

The quantum analog of a phase-space probability distribution is a **density matrix** $\widehat{\rho}$ (also called the **density operator**). This is a Hermitian operator on the Hilbert space $\mathbb{H}$ that is positive semidefinite and of unit trace:

$$\langle\Psi|\,\widehat{\rho}\,|\Psi\rangle \geq 0, \qquad \forall\,|\Psi\rangle\,; \qquad \mathrm{Tr}\,\widehat{\rho} = 1. \tag{3.137}$$

These conditions imply that the eigenvalues ρ_n of $\widehat{\rho}$ are nonnegative and sum up to 1. The physical interpretation is that the system can be found in the eigenstate $|\Psi_n\rangle$ belonging to the eigenvalue ρ_n with a probability ρ_n. More generally, the average of the observable A associated with the operator $\widehat{A}$ is given by

$$\langle A\rangle = \mathrm{Tr}\left(\widehat{\rho}\widehat{A}\right). \tag{3.138}$$

The quantum analogs of the classical equilibrium ensembles, like the canonical and p-T ones, have expressions that closely resemble the classical ones. For example, the density matrix of the canonical distribution is defined by

$$\widehat{\rho} = \frac{1}{Z}\exp\left(-\frac{\widehat{\mathcal{H}}}{k_{\mathrm{B}}T}\right), \tag{3.139}$$

where $\widehat{\mathcal{H}}$ is the Hamiltonian operator and the normalization factor Z is given by

$$Z = \mathrm{Tr}\exp\left(-\frac{\widehat{\mathcal{H}}}{k_{\mathrm{B}}T}\right). \tag{3.140}$$

In the same vein, the quantum analog of the p-T ensemble is given by

$$\widehat{\rho} = \frac{1}{Z}\exp\left(-\frac{\widehat{\mathcal{H}} + pV}{k_{\mathrm{B}}T}\right), \tag{3.141}$$

where V is the volume of the system and where Z is defined by

$$Z = \mathrm{Tr}\exp\left(-\frac{\widehat{\mathcal{H}} + pV}{k_{\mathrm{B}}T}\right). \tag{3.142}$$

The relations between the partition function Z and the thermodynamic functions straightforwardly follow. We have, for example, in the canonical ensemble,

$$-k_{\mathrm{B}}T\ln Z = F, \tag{3.143}$$

where F is the Helmholtz free energy. To obtain this relation, observe that since $F = E - TS$ with $E = \langle \widehat{\mathcal{H}} \rangle$ and $S = -\partial F/\partial T)_V$, F satisfies the differential equation

$$\langle \widehat{\mathcal{H}} \rangle = F - T \left(\frac{\partial F}{\partial T} \right)_V . \tag{3.144}$$

Let us now define F by eq. (3.143) and use the identity

$$\mathrm{Tr} \exp \left[\beta \left(F - \widehat{\mathcal{H}} \right) \right] = 1, \tag{3.145}$$

where $\beta = 1/k_B T$. Taking the derivative with respect to β of this identity, we obtain

$$\mathrm{Tr} \left\{ \exp \left[\beta \left(F - \widehat{\mathcal{H}} \right) \right] \left(F - \widehat{\mathcal{H}} + \beta \left(\frac{\partial F}{\partial \beta} \right)_V \right) \right\} = 0, \tag{3.146}$$

which corresponds to

$$F - \langle \widehat{\mathcal{H}} \rangle + T \left(\frac{\partial F}{\partial T} \right)_V = 0. \tag{3.147}$$

Thus F satisfies the same differential equation as the Helmholtz free energy. The boundary condition is given by the limit of eq. (3.143) for $T \to 0$. One obtains the ground-state energy of $\widehat{\mathcal{H}}$, which corresponds to the value of F in the same limit. Thus F can be identified with the Helmholtz free energy.

The evaluation of thermodynamic properties for quantum systems usually requires more sophisticated methods than those we use in the present book, due to the noncommutativity of operators. We shall confine ourselves to a few simple but important cases in which the calculations can be carried out by elementary methods.

3.16 The Grand Canonical Ensemble

The generalized ensemble in which the number of particles fluctuates and the chemical potential is fixed is called the grand canonical ensemble. We can directly write

$$\langle A \rangle = \frac{1}{Z_{GC}} \sum_{N=0}^{\infty} \int \mathrm{d}x_N \, A(x) \, \exp \left[- \frac{\mathcal{H}_N(x_N) - \mu N}{k_B T} \right], \tag{3.148}$$

where $\mathrm{d}x_N$ is the element of space volume and $\mathcal{H}_N(x_N)$ is the Hamiltonian of a system with N particles. The grand canonical partition function is related to the thermodynamic potential Ω:

$$\ln Z_{GC} = - \frac{E - TS - \mu N}{k_B T} = - \frac{\Omega}{k_B T}. \tag{3.149}$$

It is useful to remind that $\Omega = -pV$ for simple fluids, as a consequence of Euler's equation (see eq. (2.125)). The averages of $\mathcal{H}$ and of N are obtained, as usual, by taking

derivatives:

$$\langle \mathcal{H} \rangle = - \left. \frac{\partial \ln Z_{\mathrm{GC}}}{\partial (1/k_{\mathrm{B}} T)} \right)_{\mu/k_{\mathrm{B}} T} ; \tag{3.150}$$

$$\langle N \rangle = \left. \frac{\partial \ln Z_{\mathrm{GC}}}{\partial (\mu/k_{\mathrm{B}} T)} \right)_{1/k_{\mathrm{B}} T} . \tag{3.151}$$

By taking further derivatives, we can obtain expressions for the variances and the correlations. The derivation is left as an exercise.

Exercise 3.7 Consider a simple fluid, described by a grand canonical ensemble. Express the following quantities in terms of the grand canonical partition function:

$$\left\langle \Delta \mathcal{H}^2 \right\rangle = \left\langle \mathcal{H}^2 \right\rangle - \langle \mathcal{H} \rangle^2 ;$$

$$\left\langle \Delta N^2 \right\rangle = \left\langle N^2 \right\rangle - \langle N \rangle^2 ;$$

$$\langle \Delta \mathcal{H} \, \Delta N \rangle = \langle \mathcal{H} \, N \rangle - \langle \mathcal{H} \rangle \, \langle N \rangle .$$

Exercise 3.8 Show that the energy fluctuations in the grand canonical ensemble are always greater than the corresponding fluctuations in the canonical ensemble:

$$\left\langle \Delta \mathcal{H}^2 \right\rangle_{\mathrm{GC}(\mu)} \geq \left\langle \Delta \mathcal{H}^2 \right\rangle_{\mathrm{C}(N(\mu))},$$

where GC(μ) refers to the grand canonical ensemble with the value μ of the chemical potential, C(N) refers to the canonical one with the value N of the number of particles, and $N(\mu)$ is the average of N with the given value of the thermodynamic potential μ.

Since one usually tries to obtain the thermodynamic properties of gases as a function of the number of particles N rather than of the thermodynamic potential μ, it is worthwhile to adopt the following strategy. One calculates the grand canonical partition function, thus obtaining $\ln Z(T, \mu, V)$. One then calculates the average number of particles $\langle N \rangle$ as a function of μ, T, and V using eq. (3.151). One inverts this relation, obtaining μ as a function of $\langle N \rangle$, T, and V. This is always possible, since stability implies that $\langle N \rangle$ is a monotonically increasing function of μ when the other quantities are kept fixed. One performs the substitutions in eqs. (3.149) and (3.150) so as to obtain the equation of state for the pressure p and the internal energy $E = \langle \mathcal{H} \rangle$.

Let us calculate the partition function of the ideal monoatomic gas in the grand canonical ensemble:

$$\begin{aligned} Z_{\mathrm{GC}} &= \sum_{N=0}^{\infty} \mathrm{e}^{\mu N/k_{\mathrm{B}} T} \frac{1}{N!} \frac{1}{h^{3N}} \int \mathrm{d}\boldsymbol{p}^N \, \mathrm{d}\boldsymbol{r}^N \, \exp\left[- \frac{1}{k_{\mathrm{B}} T} \left(\frac{p_i^2}{2m} + v(\boldsymbol{r}_i) \right) \right] \\ &= \sum_{N=0}^{\infty} z^N Z_N, \end{aligned} \tag{3.152}$$

where Z_N is the canonical partition function of a gas of N particles (calculated in section 3.12.2) and where we have also introduced the **fugacity** z,

$$z = \exp\left(\frac{\mu}{k_B T}\right). \tag{3.153}$$

From eq. (3.97) we obtain

$$Z_N = \frac{V^N}{N!\lambda_B^{3N}}, \tag{3.154}$$

and therefore

$$Z_{GC} = \sum_{N=0}^{\infty} \frac{(zV)^N}{N!\lambda_B^{3N}} = \exp\left(\frac{zV}{\lambda_B^3}\right). \tag{3.155}$$

By taking the logarithm, we obtain the equation of state expressed as a function of z:

$$\frac{pV}{k_B T} = \ln Z_{GC} = \frac{zV}{\lambda_B^3}. \tag{3.156}$$

In order to express this in terms of the usual variables, let us calculate the average value of the number of particles as a function of the fugacity:

$$\langle N\rangle = \left.\frac{\partial \ln Z_{GC}}{\partial(\mu/k_B T)}\right)_{1/k_B T} = z\frac{\partial \ln Z}{\partial z} = \frac{zV}{\lambda_B^3}. \tag{3.157}$$

From this equation, we can obtain z as a function of $\langle N\rangle$ and substitute it in the preceding equation, so as to obtain the usual equation of state. Obviously, in this case it is trivial, and we immediately obtain the equation of state for ideal gases.

3.17 The Gibbs Formula for the Entropy

In this section, we want to show how in any generalized ensemble, the entropy can be expressed in terms of the probability density that defines the ensemble. In order to simplify the notations, let us assume that we are dealing with a discrete phase space.

Let us assume that our generalized ensemble is obtained by setting the value of the intensive variable f, which is conjugate to the extensive variable X. The average of the observable A is given by

$$\langle A\rangle = \frac{1}{Z}\sum_x A_x \exp\left(\frac{fX_x}{k_B T}\right), \tag{3.158}$$

where the partition function Z is given by

$$Z = \sum_x \exp\left(\frac{fX_x}{k_B T}\right), \tag{3.159}$$

and the sum runs over all microstates x of the system. As we saw, the relation between the partition function and the thermodynamic potentials is

$$\ln Z = \frac{TS + f\,\langle X\rangle}{k_B T}, \tag{3.160}$$

where we make explicit the fact that the thermodynamic value of the extensive variable X is given by its ensemble average.

The probability of the microstate x is therefore given by

$$p_x = \frac{1}{Z}\exp\left(\frac{fX_x}{k_B T}\right), \tag{3.161}$$

whose logarithm is

$$\ln p_x = \frac{fX_x}{k_B T} - \ln Z = \frac{1}{k_B T}\left(fX_x - TS - f\,\langle X\rangle\right). \tag{3.162}$$

Taking the average of both sides of this equation, we obtain

$$\langle \ln p_x\rangle = \sum_x p_x \ln p_x = \frac{1}{k_B T}\left(f\,\langle X\rangle - TS - f\,\langle X\rangle\right) = -\frac{S}{k_B}. \tag{3.163}$$

We have thus obtained **Gibbs's formula for the entropy** (cf. [Gibb02, pp. 101–104]),

$$S = -k_B \ln \sum_x p_x \ln p_x. \tag{3.164}$$

The beauty of this equation is that the right-hand side has the same form for all the generalized ensembles, and it even makes sense for an arbitrary distribution $p = (p_x)$. Indeed, apart from the proportionality constant k_B, it coincides with the **Shannon entropy** $H_S(p)$ of the probability distribution p—a measure of its uncertainty introduced by C. E. Shannon in his fundamental paper, which established information theory as a scientific discipline [Shan48]. (The fundamental concepts of information theory are briefly discussed in Appendix F.) More precisely, the knowledge of the probability distribution $p = (p_x)$ of the possible outcomes x of a random event corresponds to having a certain amount of ignorance about it. This ignorance is dissipated when we know the actual outcome of the event. Then $H_S(p)$ measures the amount of information obtained on average by knowing the value of x.

We can assign to a general phase-space distribution $P(x)$ a value of the Shannon entropy $H_S(P)$ given by

$$H_S(P) = -k \int \mathrm{d}x\, P(x) \ln P(x). \tag{3.165}$$

The problem is whether we can identify this quantity (with a suitable value of the constant k) with the physical entropy even if $P(x)$ is not an equilibrium distribution. There are at least two arguments that make such an identification debatable. On the one hand, we have to decide whether the probability distribution $P(x)$ is a property of the physical system we

are considering, or a description of our knowledge about it. In the first case, we speak of *objective* or *ontic* probabilities, and in the second case of *subjective* probabilities. If we use $P(x)$ as a proxy for evaluating time averages of observables like $A(x)$, we are considering a single system over a sufficiently long observation time $\mathcal{T}$. In this case, it makes sense to intepret $P(x)$ as an objective probability. But we are then bound to evaluate it over $\mathcal{T}$, and therefore we obtain predictions that are time-independent over the scale of the observation time. This corresponds to a state of thermodynamic equilibrium (at least on the scale of $\mathcal{T}$). If we instead consider $P(x)$ as a description of our knowledge of the system, then, by Liouville's theorem (sec. 3.4), the probability distribution $P(x, t)$ does not change locally when we follow a trajectory $x(t; x_0)$ of the system:

$$P(x(t; x_0), t) = P(x_0, t{=}0), \qquad \forall x_0. \tag{3.166}$$

As a consequence, the Shannon entropy of $P(x, t)$ remains constant:

$$-k \int \mathrm{d}\xi \; P(\xi, t) \ln P(\xi, t) = -k \int \mathrm{d}x_0 \; P(x_0, t{=}0) \ln P(x_0, t{=}0), \qquad \forall t, \tag{3.167}$$

where x_0 is given by

$$x(t; x_0) = \xi. \tag{3.168}$$

This is at variance with the expected behavior of the thermodynamic entropy, which, as we know, should in general increase with the passage of time.

Let us also remark that in the usual settings dealt with in information theory, it is in principle possible to sample the probability distribution p and therefore provide a reliable estimate of H_S. In statistical mechanics, the phase space is simply too vast to allow us to do this even for the simplest system. We must therefore estimate entropy by resorting to some hypotheses about the independence of degrees of freedom, and so on, that are valid only if the state we want to describe is close to the thermodynamic equilibrium.

The debate over the use of H_S in statistical mechanics is far from settled. We give some hints about it in the recommended readings at the end of the chapter.

3.18 Variational Derivation of the Ensembles

We can use the Gibbs formula with the necessary care, being careful that we consider only thermodynamic equilibrium and states close to it. In this section, we show how the generalized ensembles can be derived from a variational principle over the functional defined by the Gibbs formula. We assume for simplicity that the microstates x are discrete.

Let us define the **Gibbs functional** $\mathcal{S}$ of the probability distribution $p = (p_x)$:

$$\mathcal{S}(p) = -k_B \sum_x p_x \ln p_x. \tag{3.169}$$

As we have seen, the value of this functional equals the thermodynamic entropy S for the generalized ensembles. Consider now probability distributions in which some extensive variables X are fixed (within a macroscopically small uncertainty) and some others

X' are allowed to fluctuate, but their averages are fixed. Then the maximal value of $\mathcal{S}$ among all such distributions is reached for the generalized ensembles, for suitable values of the intensive variables conjugate to the fluctuating extensive ones. We shall only consider here the case of discrete microstates x, but the reasoning can be extended to continuous microstates (for example, in phase space) by considering the functional

$$\mathcal{S}[P] = -k_B \int \mathrm{d}x\, P(x) \ln P(x), \tag{3.170}$$

where the integral runs over the whole phase space Γ of the system. A **functional** is a number that depends on a whole function, in our case P. It is customary to denote functionals by enclosing their arguments in brackets.

Let us now consider the case where all thermodynamic extensive variables X have fixed values, up to small uncertainties. This identifies the accessible region Γ of phase space in which these conditions are satisfied. We now show that, among all probability distributions in phase space that vanish outside Γ, the uniform distribution has the maximal value of $\mathcal{S}$. This corresponds to the microcanonical distribution.

We thus obtain a typical problem of constrained extrema. The constraint is due to the normalization condition:

$$\sum_x p_x = 1, \tag{3.171}$$

where the sum runs over all microstates x that belong to the accessible region Γ. By introducing the Lagrange multiplier α, we look for the unconstrained extremum of the functional

$$\mathfrak{S} = \mathcal{S} - \alpha \left(\sum_x p_x - 1 \right). \tag{3.172}$$

We obtain

$$\frac{\partial \mathfrak{S}}{\partial p_x} = -k_B (\ln p_x + 1) - \alpha = 0. \tag{3.173}$$

The solution is obviously

$$p_x = \text{const.} = \frac{1}{|\Gamma|}, \quad \forall x \in \Gamma; \qquad p_x = 0, \quad \text{otherwise}, \tag{3.174}$$

where $|\Gamma|$ is the number of microstates in the accessible region. The value of the constant follows from the normalization condition. We then have

$$\mathcal{S} = k_B \ln |\Gamma| = S, \tag{3.175}$$

where S is the thermodynamical entropy, by the fundamental postulate.

We now consider the generalized ensemble defined by the intensive variable f. We then allow X to fluctuate, but add the constraint that the *average* of its conjugate variable X has a fixed value:

$$\langle X \rangle = \sum_x p_x X_x = \text{const.} \tag{3.176}$$

Let us introduce the Lagrange multiplier β and look for the extremal of

$$\mathfrak{S} = \mathcal{S} - \alpha \left(\sum_x p_x \right) - \beta \left(\sum_x p_x X_x \right). \tag{3.177}$$

(We omitted some inessential additive constants.) We obtain

$$\ln p_x = \text{const.} - \beta X_x, \tag{3.178}$$

which corresponds to the distributions of the generalized ensembles. In the case where X is the internal energy, we obtain

$$p_x \propto \mathrm{e}^{-\beta \mathcal{H}_x}, \tag{3.179}$$

where $\beta = 1/k_\mathrm{B} T$. This gives us the Boltzmann factor.

Exercise 3.9 Consider a system described by the Hamiltonian $\mathcal{H}(x)$. Show that the quantity

$$\mathcal{S}(p) = -\sum_x p(x) \ln p(x)$$

is largest for the canonical distribution

$$p^{\mathrm{eq}}(x) = \frac{\mathrm{e}^{-\mathcal{H}(x)/k_\mathrm{B} T}}{Z},$$

among all distributions $p(x)$ which have the same value of the average of $\mathcal{H}$,

$$\langle \mathcal{H} \rangle = \sum_x p(x)\, \mathcal{H}(x).$$

The cases of the p-T and the grand canonical ensembles are treated in the same way.

In the subjective probability interpetation, where $\mathcal{S}$ is identifed with the Shannon entropy H_S, the variational principle for the ensembles implies that $\mathcal{S} \leq S$ for all distributions. In other words, the equilibrium distribution (identified by whatever thermodynamic constraints we impose on the system) is the one with with the largest value of the uncertainty among all distributions that satisfy the constraints, that is, the *least informative* one among them. This approach is called the **maximum entropy** (or **MaxEnt**) **principle**.

3.19 Fluctuations of Uncorrelated Particles

In this section, we attempt to explain the universality of the equation of state of ideal gases by showing that it follows from some very simple hypotheses about the particle distribution.

Let us suppose that the phase space can be considered a product of single-particle states and that the microstate of the whole system can be defined by giving the **occupation number** for each single-particle state. For example, in the ideal quantum gas, the microstate

of the whole system is determined when the number of particles that possess a specific (quantized) value of momentum has been determined. Let us additionally assume that the occupation numbers n_x of the single-particle states x are independent random variables and that their averages are small (so that for every state x, it is possible to disregard $\langle n_x\rangle^2$ with respect to $\langle n_x\rangle$). From these hypotheses, it follows that the system obeys the equation of state of ideal gases.

Let us emphasize that nothing is said in these hypotheses about the nature of single-particle states—we can assume that they correspond to larger or smaller cells in configuration space, or in phase space, or in momentum space. One cannot even say with absolute certainty that the occupation numbers define the microstate of the system *univocally*—there is nothing to prevent different microstates from corresponding to the same occupation numbers, and so on.

Let us consider the grand canonical ensemble and express the total number of particles by means of the n_x:

$$N = \sum_x n_x. \tag{3.180}$$

We can now calculate the variance of N:

$$\left\langle N^2\right\rangle - \langle N\rangle^2 = \sum_{xx'} \left(\langle n_x n_{x'}\rangle - \langle n_x\rangle \langle n_{x'}\rangle\right) = \sum_x \left(\left\langle n_x^2\right\rangle - \langle n_x\rangle^2\right). \tag{3.181}$$

In this sum, all the terms with $x \neq x'$ vanish because we assumed the independence of the n_x for different single-particle states x. On the other hand, if $\langle n_x\rangle$ is small, the greatest contributions come from the values 0 and 1 of n_x, for which $n_x^2 = n_x$. We thus obtain

$$\begin{aligned}\left\langle N^2\right\rangle - \langle N\rangle^2 &= \sum_x \left(\left\langle n_x^2\right\rangle - \langle n_x\rangle^2\right) \simeq \sum_x \left(\langle n_x\rangle - \langle n_x\rangle^2\right) \\ &\simeq \sum_x \langle n_x\rangle = \langle N\rangle\,,\end{aligned} \tag{3.182}$$

where we have neglected $\langle n_x\rangle^2$ with respect to $\langle n_x\rangle$. The variance of N is therefore equal to its average. (We recognize here a property of the Poisson distribution.)

This relation implies the state equation for ideal gases. We have in fact

$$\langle N\rangle = \left\langle N^2\right\rangle - \langle N\rangle^2 = \frac{\partial^2 \ln Z_{\mathrm{GC}}}{\partial(\mu/k_{\mathrm{B}}T)^2}\bigg)_{T,V} = k_{\mathrm{B}}T\,\frac{\partial\,\langle N\rangle}{\partial\mu}\bigg)_{T,V}. \tag{3.183}$$

On the other hand, since μ is an intensive variable, a function of T and of the extensive thermodynamic variables V and N, one has the Euler equation

$$N\,\frac{\partial\mu}{\partial N}\bigg)_{T,V} + V\,\frac{\partial\mu}{\partial V}\bigg)_{T,N} = 0. \tag{3.184}$$

Thus from eqs. (3.183 and 3.184) we obtain

$$k_{\mathrm{B}}T = N\,\frac{\partial\mu}{\partial N}\bigg)_{T,V} = -V\,\frac{\partial\mu}{\partial V}\bigg)_{T,N} = V\,\frac{\partial p}{\partial N}\bigg)_{T,V}, \tag{3.185}$$

where we have exploited a Maxwell relation. Integrating this equation with respect to N, with the obvious boundary condition $p(N{=}0) = 0$ yields

$$pV = N k_{\rm B} T. \tag{3.186}$$

Recommended Reading

I have followed the point of view laid forth in S.-K. Ma, *Statistical Mechanics* (Singapore: World Scientific, 1985), which is close to that followed in L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part I*, 3rd ed., Oxford, UK: Pergamon Press, 1980. The maximum entropy approach is advocated, among others, by R. Balian, *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics*, Berlin/New York: Springer, 1991, chap. 4, and E. T. Jaynes, *Probability Theory: The Logic of Science* (edited by G. L. Bretthorst), Cambridge, UK: Cambridge University Press, 2003, chap. 11. Boltzmann's conception of entropy is discussed in J. L. Lebowitz, "Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy," *Physica A* **194** 1–27 (1993), where it is contrasted with the ensemble definition due to Gibbs. A recent discussion (at a more advanced level) is contained in G. Goldstein et al., "Gibbs and Boltzmann entropy in classical and quantum mechanics," in V. Allori (ed.), *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, Singapore: World Scientific, 2020.

In this book, I have chosen not to discuss the many subtle problems posed by the foundations of statistical mechanics. An interesting discussion can be found in chapter 26 of Ma's book, *Statistical Mechanics*. A different point of view is expressed in G. Gallavotti, *Statistical Mechanics: A Short Treatise*, Berlin: Springer, 1999, especially in chap. III.

4 Interaction-Free Systems

> The simplest possible variable is one that can have two values. (If there is only one value, no variation is possible.)
>
> —Shang-Keng Ma

In this chapter, we will study some simple systems in which the partition function can be expressed in terms of integrals over a small number of variables. This is possible when, in some representation, the Hamiltonian can be written as a sum of independent functions, where each one depends on a small number of arguments. In this case, by means of a suitable choice of ensemble, the partition function factorizes into many independent integrals that can be calculated explicitly. We have already found this situation in the case of the ideal gas and of the Ising paramagnet.

4.1 Harmonic Oscillators

In this section, we consider a system composed of N one-dimensional independent harmonic oscillators. The microstate of the system is defined by the collection of all r_i coordinates (which are defined so as to vanish at the equilibrium point) and their conjugate momenta p_i. The Hamiltonian has the following expression:

$$\mathcal{H}(x) = \sum_{i=1}^{N} \left(\frac{p_i^2}{2m} + \frac{1}{2}\kappa r_i^2 \right), \tag{4.1}$$

where $x = ((p_1, r_1), \ldots, (p_N, r_N))$ is the phase-space point, m is the particles' mass, and κ is Hooke's constant, connected to the angular frequency of oscillations ω_0 by

$$\omega_0^2 = \frac{\kappa}{m}. \tag{4.2}$$

The partition function is given by:

$$Z = \int \prod_{i=1}^{N} \left(\frac{\mathrm{d}p_i\, \mathrm{d}r_i}{h} \right) \exp\left[-\frac{1}{k_B T} \sum_{i=1}^{N} \left(\frac{p_i^2}{2m} + \frac{1}{2}\kappa r_i^2 \right) \right]. \tag{4.3}$$

The exponential factorizes into the product of factors, each dependent on a single pair (p_i, r_i). We thus obtain

$$\begin{aligned} Z &= \prod_{i=1}^{N} \left\{ \int \frac{\mathrm{d}p_i\, \mathrm{d}r_i}{h} \exp\left[-\frac{1}{k_\mathrm{B} T} \left(\frac{p_i^2}{2m} + \frac{1}{2} m\omega_0^2 r_i^2 \right) \right] \right\} \\ &= \left[\left(\frac{2\pi m k_\mathrm{B} T}{h^2} \right)^{1/2} \left(\frac{2\pi k_\mathrm{B} T}{m\omega_0^2} \right)^{1/2} \right]^N = \left(\frac{k_\mathrm{B} T}{\hbar\omega_0} \right)^N . \end{aligned} \tag{4.4}$$

We can thus easily obtain the internal energy and the specific heat:

$$E = -\frac{\partial \ln Z}{\partial (1/k_\mathrm{B} T)} = N k_\mathrm{B} T; \tag{4.5}$$

$$C = \frac{\partial E}{\partial T} = N k_\mathrm{B}. \tag{4.6}$$

Exercise 4.1 (Kappler's experiment) In his 1931 thesis [Kapp31], E. Kappler measured the angular fluctuations of a mirror hanging on a thin vertical wire fixed above and below, and immersed in a room at the temperature $T = 287.1$ K. He carefully measured the angular Hooke constant of the wire, obtaining $\kappa = 9.428 \cdot 10^{-9}\,\mathrm{g\,cm^2\,s^{-2}}$. He obtained a value of the variance of the angular deviation from equilibrium equal to $\langle \varphi^2 \rangle = 4.178 \cdot 10^{-6}\,\mathrm{rad^2}$. Calculate the value of k_B starting from these data.

4.1.1 The Equipartition Theorem

It is interesting to note the "universal" aspect of these results relative to E and C—each specific parameter of the system, like the angular frequency ω_0, has disappeared. This is a special case of a more general result, called the **equipartition theorem**: when the Hamiltonian of a system depends on a degree of freedom (coordinate or momentum) via only a quadratic term, the average of this term in the canonical ensemble is given by $k_\mathrm{B} T/2$.

Let us assume that we can divide the N degrees of freedom (positions and momenta) of a classical hamiltonian $\mathcal{H}$ into two classes: $(x_1, \ldots, x_r)$ and $(x_{r+1}, \ldots, x_N)$, and write $\mathcal{H}$ as the sum of two terms:

$$\mathcal{H} = \mathcal{H}_0(x_1, \ldots, x_r, x_{r+1}, \ldots, x_N) + \mathcal{H}_1(x_{r+1}, \ldots, x_N). \tag{4.7}$$

Let us moreover assume that $\mathcal{H}_0$ is quadratic in $(x_1, \ldots, x_r)$ and is positive definite, while $\mathcal{H}_1$ does not depend on them. In these hypotheses, $\mathcal{H}_0$ satisfies Euler's equation:

$$\sum_{i=1}^{r} x_i \frac{\partial \mathcal{H}}{\partial x_i} = \sum_{i=1}^{r} x_i \frac{\partial \mathcal{H}_0}{\partial x_i} = 2\mathcal{H}_0. \tag{4.8}$$

Then, the average of $\mathcal{H}_0$ in the canonical ensemble is given by

$$\langle \mathcal{H}_0 \rangle = \frac{r}{2} k_\mathrm{B} T. \tag{4.9}$$

Indeed, introducing the partition function $Z = \int \prod_{i=1}^{N} \mathrm{d}x_i \, \exp(-\mathcal{H}/k_\mathrm{B}T)$, the average has the expression

$$\begin{aligned}
\langle \mathcal{H}_0 \rangle &= \frac{1}{Z} \int \prod_{i=1}^{N} \mathrm{d}x_i \int \prod_{j=r+1}^{N} \mathrm{d}x_j \, \mathcal{H}_0 \, \exp\left(-\frac{\mathcal{H}_0 + \mathcal{H}_1}{k_\mathrm{B}T}\right) \\
&= \frac{1}{Z} \int \prod_{i=1}^{N} \mathrm{d}x_i \int \prod_{j=r+1}^{N} \mathrm{d}x_j \left(\frac{1}{2} \sum_{i=1}^{r} x_i \frac{\partial \mathcal{H}_0}{\partial x_i}\right) \exp\left(-\frac{\mathcal{H}_0 + \mathcal{H}_1}{k_\mathrm{B}T}\right) \\
&= -\frac{k_\mathrm{B}T}{2} \sum_{i=1}^{r} \left[\frac{1}{Z} \int \prod_{i=1}^{N} \mathrm{d}x_i \int \prod_{j=r+1}^{N} \mathrm{d}x_j \, x_i \frac{\partial}{\partial x_i} \exp\left(-\frac{\mathcal{H}_0 + \mathcal{H}_1}{k_\mathrm{B}T}\right) \right].
\end{aligned} \tag{4.10}$$

In this fashion, we obtain r terms, each of which can be integrated by parts. Each of them contains an integral of the form

$$\begin{aligned}
&-\frac{k_\mathrm{B}T}{2} \frac{1}{Z} \int \mathrm{d}x_i \, x_i \frac{\partial}{\partial x_i} \exp\left(-\frac{\mathcal{H}_0 + \mathcal{H}_1}{k_\mathrm{B}T}\right) \\
&\quad = -\frac{k_\mathrm{B}T}{2} \frac{1}{Z} \left[x_i \exp\left(-\frac{\mathcal{H}_0 + \mathcal{H}_1}{k_\mathrm{B}T}\right) \right]_{x_i=-\infty}^{x_i=+\infty} + \frac{k_\mathrm{B}T}{2} \frac{1}{Z} \int \mathrm{d}x_i \, \exp\left(-\frac{\mathcal{H}_0 + \mathcal{H}_1}{k_\mathrm{B}T}\right).
\end{aligned} \tag{4.11}$$

The first term vanishes by the hypothesis that the quadratic form $\mathcal{H}_0$ is positive definite. The second term can be integrated over the remaining degrees of freedom to yield Z, which cancels with the Z in the denominator. We thus obtain $k_\mathrm{B}T/2$ for each degree of freedom.

The importance of this theorem lies in the fact that any classical stable Hamiltonian is quadratic in the momenta, and "almost quadratic" in the deviations of the coordinates from the point of equilibrium (in other words, from the minimum of the potential energy). Therefore, if the internal energy is small enough (in other words, for small temperature values) we expect that it will be approximathely equal to $k_\mathrm{B}T/2$ per degree of freedom, up to an additive constant, and that the specific heat is therefore equal to $k_\mathrm{B}/2$ per degree of freedom. The results obtained for the ideal gas and the harmonic oscillators are only special cases of this result.

4.1.2 Quantum Harmonic Oscillators

The situation changes if we consider quantum systems. We know from quantum mechanics that it is possible to specify a base $|n\rangle$ in the Hilbert space of the harmonic oscillator, in which each state $|n\rangle$ is an eigenstate of the Hamiltonian, belonging to the eigenvalue $\epsilon_n = (n + 1/2)\hbar\omega_0$, $n = 0, 1, 2, \ldots$. Since we are dealing solely with the Hamiltonian operator, we can easily evaluate the partition function by weighting these states with the Boltzmann factor $\exp(-\epsilon_n/k_\mathrm{B}T)$. The situation would be different if we had to deal with several operators that do not mutually commute.

Let us consider a system of N identical and independent quantum harmonic oscillators. In this case, the partition function can be expressed as follows:

$$\begin{aligned} Z &= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\cdots\sum_{n_N=0}^{\infty}\exp\left\{-\frac{1}{k_{\mathrm{B}}T}\sum_{i=1}^{N}\left[\left(n+\frac{1}{2}\right)\hbar\omega_0\right]\right\} \\ &= \left\{\sum_{n=0}^{\infty}\exp\left[-\frac{1}{k_{\mathrm{B}}T}\left(n+\frac{1}{2}\right)\hbar\omega_0\right]\right\}^N = \zeta^N. \end{aligned} \tag{4.12}$$

Apart from a factor, ζ is a geometric series:

$$\begin{aligned} \zeta &= \exp\left(-\frac{\hbar\omega_0}{2k_{\mathrm{B}}T}\right)\sum_{n=0}^{\infty}\left[\exp\left(-\frac{\hbar\omega_0}{k_{\mathrm{B}}T}\right)\right]^n \\ &= \exp\left(-\frac{\hbar\omega_0}{2k_{\mathrm{B}}T}\right)\left[1-\exp\left(-\frac{\hbar\omega_0}{k_{\mathrm{B}}T}\right)\right]^{-1}. \end{aligned} \tag{4.13}$$

Therefore, introducing the variable $\beta = 1/k_{\mathrm{B}}T$ as in sec. 3.2.1,

$$\begin{aligned} E &= -\left.\frac{\partial \ln Z}{\partial\beta}\right|_{\beta=1/k_{\mathrm{B}}T} = -N\left.\frac{\partial \ln \zeta}{\partial\beta}\right|_{\beta=1/k_{\mathrm{B}}T} \\ &= N\frac{\hbar\omega_0}{2} + N\hbar\omega_0\exp\left(-\frac{\hbar\omega_0}{k_{\mathrm{B}}T}\right)\left[1-\exp\left(-\frac{\hbar\omega_0}{k_{\mathrm{B}}T}\right)\right]^{-1}. \end{aligned} \tag{4.14}$$

The first term represents the **zero-point energy**—in other words, the energy the harmonic oscillators possess even in their ground states as a result of the uncertainty principle. The second term represents the dependence on temperature. Taking the derivative with respect to T, we obtain the specific heat:

$$C = \frac{\partial E}{\partial T} = Nk_{\mathrm{B}}\left(\frac{\hbar\omega_0}{k_{\mathrm{B}}T}\right)^2\exp\left(-\frac{\hbar\omega_0}{k_{\mathrm{B}}T}\right)\left[1-\exp\left(-\frac{\hbar\omega_0}{k_{\mathrm{B}}T}\right)\right]^{-2}. \tag{4.15}$$

In figure 4.1, the specific heat per oscillator (expressed in k_{B} units) is shown as a function of the ratio $k_{\mathrm{B}}T/\hbar\omega_0$. We see that while for $T\to\infty$ one recovers the classical value $C = Nk_{\mathrm{B}}$, the specific heat rapidly tends to zero as soon as the temperature falls below a characteristic temperature $T_{\mathrm{E}} = \hbar\omega_0/k_{\mathrm{B}}$, called the **Einstein temperature**.

The first quantum model for the specific heat of solids was introduced by Einstein along the lines of the present calculation [Eins07]. Einstein was attempting to explain the fact that in certain solids, such as diamond, the specific heat at room temperature is smaller than the classical value $3Nk_{\mathrm{B}}$—equal to approximately 6 calories per mole (the **Dulong-Petit law**). He modelled a solid by assuming that each atom in it would oscillate about its equilibrium position with a given frequency ω_0, and, following Planck's ideas developed in the context of black body radiation, by also supposing that the values of the energy of these oscillators would be quantized in units of $\hbar\omega_0$. As a result, the specific heat becomes much smaller than Nk_{B} when the temperature falls below the Einstein temperature. Since $\omega_0^2 = \kappa/m$, the Einstein temperature is higher for harder and lighter solids, which are thus expected to have smaller values of the specific heat at a given temperature.

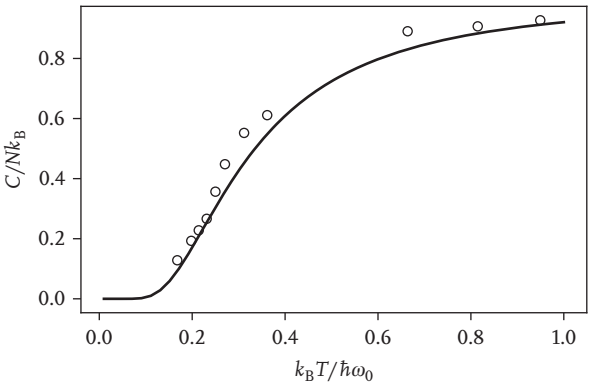


FIGURE 4.1. Specific heat per oscillator in the Einstein model, in k_B units, as a function of $k_B T/\hbar\omega_0$. Points represent Weber's data for diamond, as cited in [Eins07].

4.2 Photons and Phonons

In the present section we describe the statistical mechanics of spatially distributed oscillation modes. We first discuss the oscillations of the electromagnetic field, starting from the important case of the black body radiation. We then discuss the collective oscillations in solids, in order to go beyond the Einstein model.

4.2.1 Black Body Radiation: Photons

A physical realization of a perfectly black body is a cavity in a metal body filled with electromagnetic radiation, in equilibrium with the body at a given temperature T. In the present section, we will derive Planck's law of radiation for such a system, by building on the analogy with a system of quantum oscillators. Let us consider a cubic cavity of side length L: its volume V is therefore equal to L^3. To simplify the discussion, we impose periodic boundary conditions. Therefore, the oscillations of the electromagnetic field in the cavity must be spatially periodic, with period equal to L in each direction. As a consequence, these oscillations can be represented by a Fourier series, as a combination of elementary oscillations, each characterized by a wave vector of the form $\boldsymbol{k} = (2\pi/L)(\ell_x, \ell_y, \ell_z)$, where the ℓs are integers. (Fourier transforms are briefly reviewed in appendix D.) In addition, each oscillation is characterized by a polarization state: we choose as base states those of circular polarization, which are identified by a scalar quantity $\alpha = \pm 1$, according to whether the field rotates counterclockwise (+) or clockwise (−) as seen from $\boldsymbol{k}$. Thus $(\boldsymbol{k}, \alpha)$ identifies an oscillation mode, with a characteristic (angular) frequency $\omega(\boldsymbol{k}, \alpha) = c\,|\boldsymbol{k}|$, where c is the speed of light.

The microstate of the system is identified by assigning the eigenstate $|n(\boldsymbol{k}, \alpha)\rangle$ of each oscillating mode, where $n(\boldsymbol{k}, \alpha) \in \{0, 1, 2, \ldots\}$. It is costumary to refer to $n(\boldsymbol{k}, \alpha)$ as to the **occupation number** of the mode $(\boldsymbol{k}, \alpha)$, identified by its wave vector $\boldsymbol{k}$ and its polarization state α. We thus denote the microstate by $x = (n(\boldsymbol{k}, \alpha))$. The energy of this microstate is given by

$$\mathcal{H}(x) = \sum_{\boldsymbol{k},\alpha} \left[n(\boldsymbol{k},\alpha) + \frac{1}{2} \right] \hbar\omega(\boldsymbol{k},\alpha). \tag{4.16}$$

We can now calculate the partition function in the canonical ensemble:

$$Z = \sum_{(n(\boldsymbol{k},\alpha))} \exp\left\{ -\frac{1}{k_\mathrm{B} T} \sum_{\boldsymbol{k},\alpha} \left[\left(n(\boldsymbol{k},\alpha) + \frac{1}{2} \right) \hbar\omega(\boldsymbol{k},\alpha) \right] \right\}. \tag{4.17}$$

As in the preceding section, we can break up this sum into factors, in each of which only a single oscillation mode $(\boldsymbol{k},\alpha)$ appears—the resulting sums are geometric series that can be easily summed. We thus obtain the following expression of $\ln Z$:

$$\ln Z = \sum_{\boldsymbol{k},\alpha} \left\{ -\frac{\hbar\omega(\boldsymbol{k},\alpha)}{2k_\mathrm{B} T} - \ln\left[1 - \exp\left(-\frac{\hbar\omega(\boldsymbol{k},\alpha)}{k_\mathrm{B} T} \right) \right] \right\}. \tag{4.18}$$

In order to calculate this sum, we regroup those terms that correspond to values close to the frequency ω:

$$\ln Z = \int_0^\infty \mathrm{d}\omega \sum_{\boldsymbol{k},\alpha} \delta\left(\omega - \omega(\boldsymbol{k},\alpha)\right) \left\{ -\frac{\hbar\omega}{2k_\mathrm{B} T} - \ln\left[1 - \exp\left(-\frac{\hbar\omega}{k_\mathrm{B} T} \right) \right] \right\}. \tag{4.19}$$

In wave-vector space, there are two states with different polarizations for each small cube of volume $(2\pi/L)^3 = (2\pi)^3/V$. Making use of the relation between ω and $\boldsymbol{k}$, we obtain

$$\sum_{\boldsymbol{k},\alpha} \delta\left(\omega - \omega(\boldsymbol{k},\alpha)\right) = 2\frac{V}{(2\pi)^3} \int \mathrm{d}k\, 4\pi k^2\, \delta(\omega - ck) = 2\frac{V}{(2\pi c)^3} 4\pi\omega^2. \tag{4.20}$$

By substituting in the preceding equation, we obtain

$$\ln Z = \frac{V}{\pi^2 c^3} \int_0^\infty \mathrm{d}\omega\, \omega^2 \left\{ -\frac{\hbar\omega}{2k_\mathrm{B} T} - \ln\left[1 - \exp\left(-\frac{\hbar\omega}{k_\mathrm{B} T} \right) \right] \right\}. \tag{4.21}$$

The first term corresponds to the zero-point energy. Since the number of oscillator modes is infinite, it is infinite. However, it does not concern us for the moment. (It is, however, at the origin of the so-called Casimir effect—in other words, of the attractive forces that arise among the conductive walls delimiting a cavity.) Indeed, since the energy is defined up to an additive constant, we can choose as energy zero the internal energy of the system in the limit $T \to 0$. If we do this, we take into account only the second term. We thus obtain

$$\begin{aligned} E &= \frac{V}{\pi^2 c^3} \int_0^\infty \mathrm{d}\omega\, \omega^3 \exp\left(-\frac{\hbar\omega}{k_\mathrm{B} T} \right) \left[1 - \exp\left(-\frac{\hbar\omega}{k_\mathrm{B} T} \right) \right]^{-1} \\ &= \frac{V}{\pi^2 \hbar^4 c^3} (k_\mathrm{B} T)^4 \int_0^\infty \mathrm{d}x\, \frac{x^3}{\mathrm{e}^x - 1}. \end{aligned} \tag{4.22}$$

We have

$$\begin{aligned} \int_0^\infty \mathrm{d}x\, \frac{x^3}{\mathrm{e}^x - 1} &= \int_0^\infty \mathrm{d}x\, x^3 \sum_{k=1}^\infty \mathrm{e}^{-kx} = \sum_{k=1}^\infty \frac{1}{k^4} \int_0^\infty \mathrm{d}x\, x^3\, \mathrm{e}^{-x} \\ &= \zeta_\mathrm{R}(4)\, \Gamma_\mathrm{E}(4), \end{aligned} \tag{4.23}$$

where $\Gamma_E(s)$ is Euler's gamma function and $\zeta_R(s)$ is the Riemann zeta function, defined by

$$\zeta_R(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}. \tag{4.24}$$

More specifically, we have $\Gamma_E(4) = 3! = 6$ and

$$\zeta_R(4) = \frac{\pi^4}{90}. \tag{4.25}$$

We thus obtain

$$E = \frac{\pi}{15} \frac{V}{\hbar^4 c^3} (k_B T)^4. \tag{4.26}$$

By taking the derivative of the second term in eq. (4.18) with respect to $1/k_B T$, we obtain a more transparent expression of the internal energy E. We obtain, remembering that the characteristic energy of the frequency mode $(\boldsymbol{k}, \alpha)$ is equal to $\hbar\omega(\boldsymbol{k}, \alpha)$:

$$\begin{aligned} E &= \frac{V}{(2\pi)^3} \sum_{\alpha} \int \mathrm{d}\boldsymbol{k}\, \hbar\omega(\boldsymbol{k}, \alpha) \left[\exp\left(\frac{\hbar\omega(\boldsymbol{k}, \alpha)}{k_B T} - 1 \right) \right]^{-1} \\ &= \frac{V}{(2\pi)^3} \sum_{\alpha} \int \mathrm{d}\boldsymbol{k}\, \langle \epsilon(\boldsymbol{k}, \alpha) \rangle, \end{aligned} \tag{4.27}$$

where the average energy $\langle \epsilon(\boldsymbol{k}, \alpha) \rangle$ is given, for each mode $(\boldsymbol{k}, \alpha)$, by

$$\langle \epsilon(\boldsymbol{k}, \alpha) \rangle = \hbar\omega(\boldsymbol{k}, \alpha) \langle n(\boldsymbol{k}, \alpha) \rangle. \tag{4.28}$$

We can see from this expression that for each oscillator mode the average occupation number $\langle n(\boldsymbol{k}, \alpha) \rangle$ is given by $f_-(\hbar\omega(\boldsymbol{k}, \alpha)/k_B T)$, where

$$f_-(x) = \frac{1}{\mathrm{e}^x - 1} \tag{4.29}$$

is the **Bose factor**. Its behavior is shown in figure 4.2. It diverges as $1/x$ for $x \to 0$, decreases monotonically with x and behaves as e^{-x} for $x \gg 1$.

The expression for $\langle \epsilon(\boldsymbol{k}, \alpha) \rangle$ admits two classical limits. When $\hbar\omega \ll k_B T$, it approaches $k_B T$ in accordance with the equipartition theorem (each mode behaves like a classical harmonic oscillator). When $\hbar\omega \gg k_B T$, the probability of having one energy quantum $\hbar\omega$ is proportional to the Boltzmann factor $\exp(-\hbar\omega/k_B T)$. The quanta then behave like particles. This last observation was made by Einstein on the basis of the black body radiation law and led him to introduce the concept of the **photon**.

It is useful to calculate the power $W(\omega)\,\mathrm{d}\omega$ of the radiation emitted within a certain frequency interval. One must keep in mind that this radiation escapes the cavity, for example, through a small hole, at the speed of light c. To evaluate the power emitted by the cavity through a hole of area S, in a frequency interval of size $\mathrm{d}\omega$, we must calculate the energy density $\mathcal{E}(\mathrm{d}\omega)$ per unit volume of all the oscillation modes whose frequency lies in this interval, and whose wave vectors are directed toward the exterior of the cavity. The wave

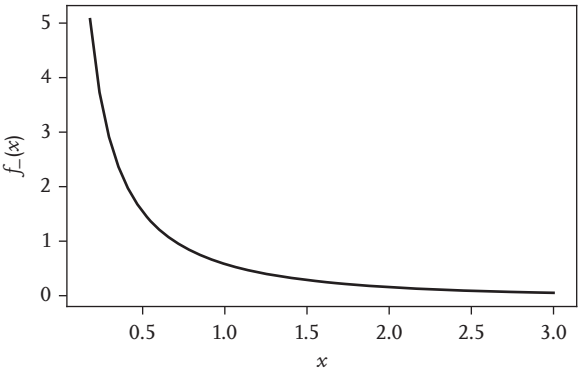


FIGURE 4.2. Bose factor: $f_-(x) = (e^x - 1)^{-1}$.

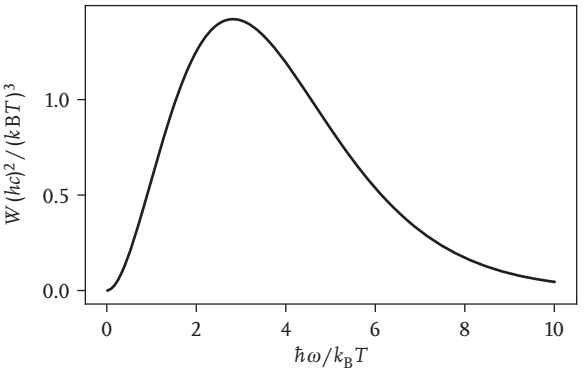


FIGURE 4.3. Planck's law for the black body radiation.

vector $\boldsymbol{k}$ of these modes satisfies the relation $\omega \leq c\,|\boldsymbol{k}| \leq \omega + d\omega$. Summing over the polarization modes α and over the directions of $\boldsymbol{k}$ toward the exterior, we obtain that the power emitted per unit frequency and per unit area of the hole is given by

$$W(\omega) = \frac{\hbar}{4\pi^2 c^2} \frac{\omega^3}{\exp(\hbar\omega/k_B T) - 1}, \tag{4.30}$$

which is Planck's formula for the black body radiation [Plan01].

In figure 4.3, the intensity W of the black body radiation is plotted (in units $(k_B T)^3/(hc)^2$) as a function of the frequency ω (in units $k_B T/\hbar$). One should note the classical regime at low frequencies ($\hbar\omega \ll k_B T$), where $W \sim \omega^2$, and the "photonic" regime at high frequencies, where $W \sim e^{-\hbar\omega/k_B T}$, known as **Wien's law**.

Notice that Planck's law has only one parameter: the temperature. Therefore, one can evaluate the temperature of a radiating body by measuring the intensity emitted at a single wavelength. This allowed Penzias and Wilson, in their celebrated work of 1965, to evaluate at 3.5 K the temperature of the residual background radiation at a wavelength of 7.35 cm. This radiation was interpreted as cosmic black body radiation, a footprint of the Big Bang (see [Penz65] and [Dick65]). The accepted value of its temperature is now about 2.725 K.

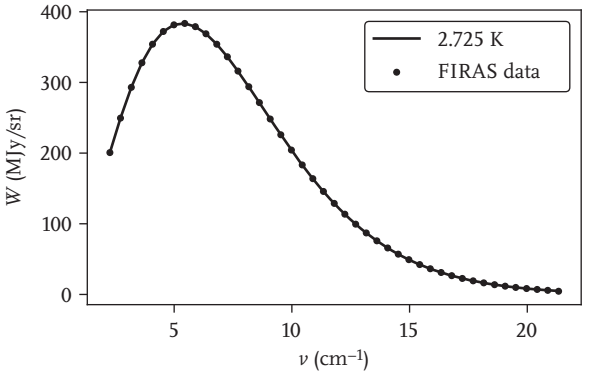


FIGURE 4.4. Spectrum of the monopole component of the cosmic background radiation compared with Planck's law at 2.725 K. The units are MJy/sr, where $\mathrm{Jy} = 10^{-26}\,\mathrm{Wm^{-2}Hz^{-1}}$ and sr is the steradian. Data from the COBE Far InfraRed Absolute Spectrophotometer [FIRAS].

Exercise 4.2 (Stefan-Boltzmann law) By integrating eq. (4.30), one obtains the total power emitted by a body at temperature T per unit surface:

$$W_{\mathrm{tot}} = \int_0^\infty \mathrm{d}\omega\; W(\omega) = \sigma_{\mathrm{S}} T^4,$$

where the **Stefan-Boltzmann constant** σ_{S} is given by

$$\sigma_{\mathrm{S}} = \frac{2\pi^5 k_{\mathrm{B}}^4}{15 h^3 c^2} \simeq 5.670 \cdot 10^{-8}\,\mathrm{W/m^2K^4}.$$

Show that this result (except for the value of the constant σ_{S}) follows from thermodynamics and the following hypotheses:

1. The internal energy of the black body radiation is proportional to the system's volume and the energy density depends only on the temperature:

   $$E = \mathcal{E}(T)V.$$

2. The pressure p is equal to 1/3 of the energy density:

   $$p = \frac{1}{3}\mathcal{E}(T).$$

   This hypothesis can be justified by taking account of the fact that if the distribution of light is isotropic, only one third of the light present contributes to the pressure on any given element of the surface.

Exercise 4.3 Evaluate the power needed to keep the human body at the constant temperature of 37°C, under the following hypotheses:

- The body behaves like a black body, with an emissivity $E = 0.7$ (fraction of the radiation emitted or received).

- The surface area is approximately $2\,\mathrm{m}^2$.
- The body receives the radiation of the environment at 27°C.

Compare it with the basal metabolism rate, estimated at $1.1 \cdot 10^7\,\mathrm{J/day}$.

4.2.2 Phonons and the Specific Heat of Solids

Einstein's model is not very efficient in describing in a quantitative way the behavior of specific heat in solids. The fundamental reason for this state of affairs was identified at the same time by Born and von Kármán (see [Born12a] and [Born12b]) on the one hand and by Debye [Deby12] on the other. Debye's model, although less rigorous, is more intuitive. The idea is that the oscillations of atoms in a crystal lattice are not independent of one another, but are collective oscillations, and can be characterized by a wave vector $\boldsymbol{k}$ and a polarization index α. Once these quantities are assigned, the oscillations are harmonic, with a frequency $\omega(\boldsymbol{k}, \alpha)$. The oscillations of the lattice can be interpreted as sound waves. The relation between the wave vector and the frequency is given, at least for small wave vectors, by

$$\omega(\boldsymbol{k}, \alpha) = c_\alpha \left|\boldsymbol{k}\right|, \tag{4.31}$$

where c_α is the speed of sound for oscillations with polarization α.

The model thus obtained for a solid at temperature T is fairly similar to that of a radiating cavity, which is filled by photons. By analogy, the quanta of the oscillation of the crystal lattice are called **phonons**. The fundamental difference with respect to the case of the radiant cavity (apart from the fact that the states of polarization are three in solids and only one in fluids, and that the speed of sound can be different in different polarizations) is that the number of degrees of freedom of oscillation of a solid composed of N atoms is finite, and equal to $3N$. It is therefore necessary to cut off the sums over the oscillatory degrees of freedom so as to remain with no more than $3N$ total degrees of freedom. From a geometrical point of view, this is equivalent to excluding the oscillations for wavelengths smaller than the lattice step.

We have thus defined the Debye model:

$$\ln Z = -\frac{V}{(2\pi h)^3} \sum_\alpha \int_0^{k_\mathrm{D}} \mathrm{d}k\, 4\pi k^2\, \ln\left[1 - \exp\left(\frac{\hbar c_\alpha k}{k_\mathrm{B} T}\right)\right], \tag{4.32}$$

where the cutoff or **Debye wave vector** k_D is implicitly defined by

$$\frac{V}{(2\pi h)^3} \int_0^{k_\mathrm{D}} \mathrm{d}k\, 4\pi k^2 = N. \tag{4.33}$$

If we assume for simplicity that the speed of sound c does not depend on polarization, we can introduce the **Debye frequency** $\omega_\mathrm{D} = ck_\mathrm{D}$ and the **Debye temperature** $k_\mathrm{B} T_\mathrm{D} = \hbar\omega_\mathrm{D}$. Making use of these quantities, one obtains the following expression for the internal energy of the solid:

$$E = 3\frac{3N(k_\mathrm{B} T)^4}{(\hbar\omega_\mathrm{D})^3} \int_0^{\hbar\omega_\mathrm{D}/k_\mathrm{B} T} \mathrm{d}x\, \frac{x^3}{\mathrm{e}^x - 1}. \tag{4.34}$$

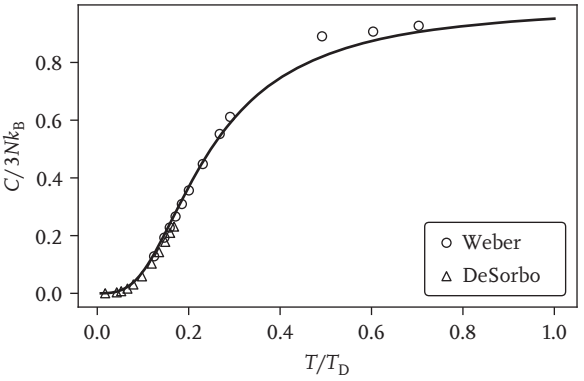


FIGURE 4.5. Specific heat per degree of freedom, $C/3Nk_B$, in the Debye model, as a function of T/T_D, where $T_D = \hbar\omega_D/k_B$ is the Debye temperature. Data for diamond: circles represent Weber's data, as cited in [Eins07], while triangles correspond to DeSorbo's low-temperature data [DeSo53].

It is convenient to introduce the **Debye function**:

$$\mathcal{D}(x) = \frac{3}{x^3}\int_0^x \mathrm{d}t\,\frac{t^3}{\mathrm{e}^t - 1}. \tag{4.35}$$

The expression (4.34) can be then rewritten as follows:

$$E = 3Nk_B T\,\mathcal{D}\left(\frac{T_D}{T}\right). \tag{4.36}$$

The expression of the specific heat then follows:

$$\begin{aligned} C &= 3Nk_B\left[\mathcal{D}\left(\frac{T_D}{T}\right) - x\left.\frac{\partial\mathcal{D}(x)}{\partial x}\right|_{x=T_D/T}\right] \\ &= 3Nk_B\left[4\mathcal{D}\left(\frac{T_D}{T}\right) - \frac{3(T_D/T)}{\mathrm{e}^{T_D/T}-1}\right]. \end{aligned} \tag{4.37}$$

The behavior of the specific heat per degree of freedom $C/3Nk_B$ in the Debye model is plotted as a function of T/T_D in figure 4.5. For high temperatures, we once again obtain the classical value $C = 3Nk_B$, while for small temperatures, we obtain

$$C \simeq 36Nk_B\left(\frac{T}{T_D}\right)^3\int_0^\infty \mathrm{d}t\,\frac{t^3}{\mathrm{e}^t-1} = Nk_B\frac{12\pi^4}{5}\left(\frac{T}{T_D}\right)^3, \tag{4.38}$$

where we neglected the second term in eq. (4.37) and exploited eq. (4.23). The specific heat is therefore proportional to T^3 at low temperatures—it vanishes much more slowly than the exponential obtained from the Einstein model. This behavior is in agreement with the experimental data for insulating materials. Indeed, the figure shows that diamond is well described by a Debye model if low-temperature data are taken into account. In metals, however, the behavior is instead linear in T, due to the contribution of the conducting electrons, as we see later.

4.2.3 Phonons and Second Sound

Since the phonons are *particles*, they behave in some sense as a gas of particles. It turns out, in particular, that the phonon gas can sustain wave-like oscillations similar in nature to those that correspond to sound in an ordinary gas. In order to understand this phenomenon, it is useful to first consider the nature of sound waves in an ordinary gas.

We consider a gas in a state described by a local density $\rho(\boldsymbol{r}, t)$ and a local velocity $\boldsymbol{v}(\boldsymbol{r}, t)$. Then the local mass density is given by $m\rho(\boldsymbol{r}, t)$ and the local momentum density is given by $m\rho(\boldsymbol{r}, t)\boldsymbol{v}(\boldsymbol{r}, t)$. Here m is the mass of a single molecule. If the gas is a mixture, it is equal to the average mass:

$$m = \frac{1}{N}\sum_i N_i m_i, \tag{4.39}$$

where the sum runs over all different molecule species i of mass m_i, where N_i is their respective number and N is the total number of molecules.

Since the number of molecules is conserved, $\rho(\boldsymbol{r}, t)$ satisfies a continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0. \tag{4.40}$$

On the other hand, by Newton's law, the total derivative of the momentum density is equal to the net force density acting on a volume element. The force density is given by $-\nabla p$, where $p(\boldsymbol{r}, t)$ denotes the pressure acting on point $\boldsymbol{r}$ at time t. To see this, consider a box of sides $\mathrm{d}x$, $\mathrm{d}y$, and $\mathrm{d}x$, and therefore of volume $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$, whose faces are normal to the coordinate axes, placed in $\boldsymbol{r}$. Then the x component of the net force acting on the box is given by $\mathrm{d}y\,\mathrm{d}z\,(p(\boldsymbol{r}) - p(\boldsymbol{r} + \boldsymbol{i}\,\mathrm{d}x)) \simeq -\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,\partial p/\partial x)_{yz}$, where $\boldsymbol{i}$ is the unit vector in the x direction. The same reasoning applies to the other directions. Using this result, we obtain

$$\frac{\partial (m\rho \boldsymbol{v})}{\partial t} + \boldsymbol{v} \cdot \nabla\,(m\rho \boldsymbol{v}) = -\nabla p. \tag{4.41}$$

We assume that the density deviates little from the equilibrium value ρ_{eq}:

$$\rho(\boldsymbol{r}, t) = \rho_{\mathrm{eq}} + \delta\rho(\boldsymbol{r}, t). \tag{4.42}$$

We also assume that the velocity $\boldsymbol{v}(\boldsymbol{r}, t)$ is small. We then expand the above equations to first order in $\delta\rho$ and $\boldsymbol{v}$. We obtain

$$\frac{\partial \rho}{\partial t} = \frac{\partial \delta\rho}{\partial t} = -\rho_{\mathrm{eq}} \nabla \cdot \boldsymbol{v}; \tag{4.43}$$

$$\rho_{\mathrm{eq}} \frac{\partial \boldsymbol{v}}{\partial t} = -\frac{1}{m}\nabla p. \tag{4.44}$$

On the other hand, we have

$$\nabla p = \frac{\partial p}{\partial \rho}\Big)_S \nabla \rho_1 = \frac{1}{\rho_{\mathrm{eq}} \kappa_S} \nabla \rho_1, \tag{4.45}$$

where κ_S is the adiabatic compressibility:

$$\kappa_S = -\frac{1}{V}\left.\frac{\partial V}{\partial p}\right)_S = \frac{1}{\rho}\left.\frac{\partial \rho}{\partial p}\right)_S. \tag{4.46}$$

The *adiabatic* (rather than the isothermal) compressibility appears here, because the system does not have time to equilibrate its temperature during an oscillation. Substituting eq. (4.45) in eq. (4.44), and taking the time derivatives of both eqs. (4.43) and (4.44), we obtain d'Alembert's equation for the sound waves:

$$\frac{\partial^2 \rho}{\partial t^2} = c^2 \nabla^2 \rho_1, \tag{4.47}$$

where the **speed of sound** c_s is given by

$$c_s = \sqrt{\frac{1}{m\rho_{eq}\kappa_S}}. \tag{4.48}$$

For an adiabatic transformation, according to eq. (2.160), one has $p\rho^{-\gamma} = \text{const.}$, where $\gamma = C_p/C_V$. Taking the logarithmic derivative of this relation we obtain $\kappa_S = 1/\gamma p$. Therefore the speed of sound is given by

$$c_s = \sqrt{\frac{\gamma p}{m\rho_{eq}}} = \sqrt{\frac{\gamma k_B T}{m}}, \tag{4.49}$$

where we have taken advantage of the equation of state. For a monoatomic gas, we have seen in sec. 3.12.2 that $\gamma = 5/3$. We shall see later that, for a biatomic gas like air, one has $\gamma = 7/5$.

Let us now consider the possibility that the phonon gas itself can oscillate. Since the number of phonons is not conserved, it is rather the energy density $\mathcal{E}$ that oscillates. Moreover, by the conservation of energy, its density satisfies a continuity equation of the form

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \boldsymbol{j}_E = 0, \tag{4.50}$$

where $\boldsymbol{j}_E$ is the energy current density. On the other hand, the local momentum density, which we shall denote by $\boldsymbol{q}$, satisfies Newton's law:

$$\frac{\partial \boldsymbol{q}}{\partial t} + \boldsymbol{v} \cdot \nabla \boldsymbol{q} = -\nabla p. \tag{4.51}$$

As for the case of ordinary sound, the second term on the left-hand side can be neglected.

We must now find a relation among the three quantities $\mathcal{E}, \boldsymbol{j}_E$, and $\boldsymbol{q}$ and the local velocity $\boldsymbol{v}$, in order to obtain a closed equation. Since the effect of the pressure gradients is to put the medium locally in motion, we must first briefly discuss the statistical mechanics of a system in motion.

Given a system with total energy E_{tot} and total momentum $\boldsymbol{P}$, we can consider $\boldsymbol{P}$ as a new extensive variable. However, since entropy is a measure of the phase space volume, it should be independent of the state of motion of the system. It should therefore

only depend on its *internal* energy E, that is, its energy as measured in the center-of-mass reference frame.[1] We obtain therefore

$$S(E_{\text{tot}}, \boldsymbol{P}) = S\left(E_{\text{tot}} - \frac{P^2}{2M}, 0\right), \tag{4.52}$$

where M is the total mass of the system. Thus the intensive variable conjugate to the total momentum is

$$\left.\frac{\partial S}{\partial \boldsymbol{P}}\right)_{E_{\text{tot}}} = \frac{\partial S}{\partial E}\left(-\frac{\boldsymbol{P}}{M}\right) = -\frac{\boldsymbol{v}}{T}. \tag{4.53}$$

Therefore, the intensive quantity conjugate to $\boldsymbol{P}$ is the velocity of the system. Using the same reasoning employed in chapter 2, one can prove that the velocity must be the same, at equilibrium, in all parts of the system. (In the same way, one can introduce the total angular momentum $\boldsymbol{M}$ as a new extensive variable, and deduce that its conjugate intensive variable is the angular velocity, which must be uniform at equilibrium.)

We now consider a system of phonons in local equilibrium, animated by the velocity $\boldsymbol{v}$. In view of the preceding discussion, the average number of phonons of wavenumber $\boldsymbol{k}$ and polarization α in such a system is given by the Bose factor $f_-(\epsilon/k_{\text{B}}T) = 1/\left(\mathrm{e}^{\epsilon/k_{\text{B}}T} - 1\right)$, where ϵ is the phonon energy in the rest frame, given by

$$\epsilon(\boldsymbol{k}, \boldsymbol{v}) = \hbar\omega(\boldsymbol{k}, \alpha) - \hbar\boldsymbol{k}\cdot\boldsymbol{v}, \tag{4.54}$$

where we have used the fact that $\hbar\boldsymbol{k}$ is the momentum associated with a phonon of wave vector $\boldsymbol{k}$. We have $\omega(\boldsymbol{k}, \alpha) = c_\alpha \left|\boldsymbol{k}\right|$. In the following, we assume that the sound velocity c_{s} is independent of the polarization, and we denote the frequency by ω_k to lighten the notation.

We thus obtain the following expressions for $\mathcal{E}$, $\boldsymbol{q}$, and $\boldsymbol{j}_E$:

$$\mathcal{E} = \frac{3}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k} f_-\left(\frac{\hbar\omega_k - \hbar\boldsymbol{k}\cdot\boldsymbol{v}}{k_{\text{B}}T}\right)\hbar\omega_k, \tag{4.55}$$

$$\boldsymbol{q} = \frac{3}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k} f_-\left(\frac{\hbar\omega_k - \hbar\boldsymbol{k}\cdot\boldsymbol{v}}{k_{\text{B}}T}\right)\hbar\boldsymbol{k}, \tag{4.56}$$

$$\boldsymbol{j}_E = \frac{3}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k} f_-\left(\frac{\hbar\omega_k - \hbar\boldsymbol{k}\cdot\boldsymbol{v}}{k_{\text{B}}T}\right)\hbar\omega_k\, c\frac{\boldsymbol{k}}{\left|\boldsymbol{k}\right|}. \tag{4.57}$$

The factor 3 in these equations comes from the sum over the three polarization states. (Let us recall that the phonon gas behaves like a fluid, but the underlying solid does not!) The last equation is due to the fact that the energy contained in the phonon moves at the speed of sound and parallel to the wave vector $\boldsymbol{k}$. Expanding to the first order in $\boldsymbol{v}$, one obtains

$$\mathcal{E}(\boldsymbol{v}) \simeq \mathcal{E}(\boldsymbol{v} = 0). \tag{4.58}$$

[1] We consider here only *nonrelativistic* systems. The discussion of relativistic systems is more difficult and, in some aspects, still open.

Let us now expand the expression of $\boldsymbol{q}$ to first order:

$$\begin{aligned} q_i &= \frac{3}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k} f_- \left(\frac{\hbar\omega_k - \hbar\boldsymbol{k}\cdot\boldsymbol{v}}{k_\mathrm{B}T}\right)\hbar k_i \\ &\simeq -\frac{3}{(2\pi)^3}\frac{1}{k_\mathrm{B}T}\sum_j \int \mathrm{d}\boldsymbol{k}\,\hbar k_j v_j f_-'\left(\frac{\hbar\omega_k}{k_\mathrm{B}T}\right)\hbar k_i. \end{aligned} \tag{4.59}$$

For an arbitrary function $f(k)$ of $k = |\boldsymbol{k}|$ we have

$$\int \mathrm{d}\boldsymbol{k}\, k_i k_j f(k) = \delta_{ij}\int \mathrm{d}\boldsymbol{k}\, k_i^2 f(k) = \frac{1}{3}\delta_{ij}\int \mathrm{d}\boldsymbol{k}\, k^2 f(k), \tag{4.60}$$

where δ_{ij} is the **Kronecker delta**, defined by

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{4.61}$$

We obtain therefore

$$\boldsymbol{q} = -\boldsymbol{v}\frac{\hbar^2}{(2\pi)^3 k_\mathrm{B}T}\int \mathrm{d}\boldsymbol{k}\, k^2 f_-'\left(\frac{\hbar\omega_k}{k_\mathrm{B}T}\right) = \rho_\mathrm{II}\,\boldsymbol{v}, \tag{4.62}$$

where ρ_II plays the role of mass density for the phonon gas and is given by

$$\rho_\mathrm{II} = \frac{2\pi(k_\mathrm{B}T)^4}{15\hbar^3 c^5}. \tag{4.63}$$

In the same fashion, we obtain

$$\boldsymbol{j}_E \simeq \frac{3}{(2\pi)^3 k_\mathrm{B}T}\int \mathrm{d}\boldsymbol{k}\,(-\hbar\boldsymbol{k}\cdot\boldsymbol{v}) f_-'\left(\frac{\hbar\omega_k}{k_\mathrm{B}T}\right)\hbar\omega_k\, c_\mathrm{s}\frac{\boldsymbol{k}}{|\boldsymbol{k}|} = c_\mathrm{s}^2\rho_\mathrm{II}\boldsymbol{v}. \tag{4.64}$$

On the other hand, we have

$$\frac{\partial\mathcal{E}}{\partial t} = \frac{\partial\mathcal{E}}{\partial T}\frac{\partial T}{\partial t}; \tag{4.65}$$

$$\nabla p = \frac{1}{3}\nabla\mathcal{E} = \frac{1}{3}\frac{\partial\mathcal{E}}{\partial T}\nabla T. \tag{4.66}$$

The last relation comes from the relation $pV = E/3$, which holds for phonons as a result of the same isotropy reasons that make it valid for photons. We have moreover

$$\begin{aligned} \frac{\partial\mathcal{E}}{\partial T} &= \frac{3}{(2\pi)^3}\int \mathrm{d}\boldsymbol{k}\,\frac{\partial}{\partial T} f_-\left(\frac{\hbar\omega_k}{k_\mathrm{B}T}\right)\hbar\omega_k \\ &= -\frac{3}{(2\pi)^3 k_\mathrm{B}T^2}\int \mathrm{d}\boldsymbol{k}\,(\hbar\omega_k)^2 f_-'\left(\frac{\hbar\omega_k}{k_\mathrm{B}T}\right) = \frac{3c_\mathrm{s}^2}{T}\rho_\mathrm{II}. \end{aligned} \tag{4.67}$$

Substituting in eqs. (4.50) and (4.51) we obtain

$$\frac{3}{T}\frac{\partial T}{\partial t} = -\nabla\cdot\boldsymbol{v}, \tag{4.68}$$

$$\frac{\partial\boldsymbol{v}}{\partial t} = -\frac{c_\mathrm{s}^2}{T}\nabla T. \tag{4.69}$$

Taking the time derivative of eq. (4.68) and substituting eq. (4.69) in it, we finally obtain d'Alembert's equation for the second sound:

$$\frac{\partial^2 T}{\partial t^2} = c_{\mathrm{II}}^2 \nabla^2 T, \tag{4.70}$$

where we have neglected terms of higher order in ∇T, and we have introduced the **second sound speed**

$$c_{\mathrm{II}} = \frac{c_{\mathrm{s}}}{\sqrt{3}}. \tag{4.71}$$

These temperature oscillations will be observable if the hypotheses we made are satisfied. These hypotheses state that

- the variation in the phonon momentum density is due only to pressure gradients;
- the interaction within the phonon gas, as well as the interactions between phonons and the lattice, are fairly weak, to the extent that one can consider phonon gas to be an ideal gas.

These hypotheses are satisfied if the sample is fairly pure, since the interaction with impurities does not conserve momentum, and if the temperature is so low that one can disregard those interactions among phonons that do not conserve momentum. If the temperature is too low, however, the phonon gas becomes too dilute and is no more able to reach equilibrium. A thorough discussion is contained in [Proh64].

4.3 Boson and Fermion Gases

Let us consider a set of noninteracting quantum particles, enclosed in a cubic box of side L, and therefore of volume $V = L^3$. We want to study the thermodynamic properties of this system. The microstate of the system is identified by the collection of the occupation numbers of each single-particle state, since the particles are indistiguishable. Moreover, since there is no interaction, such a state is an eigenstate of the Hamiltonian, and therefore possesses a well-defined energy value.

The first problem is therefore to classify the single-particle states. If we assume that the particles do not have internal degrees of freedom, only the translational degrees of freedom and spin remain. In order to fix the translational degrees of freedom, we impose periodic boundary conditions. In this case, the single-particle wave function must satisfy the conditions

$$\psi(x + L, y, z) = \psi(x, y + L, z) = \psi(x, y, z + L) = \psi(x, y, z). \tag{4.72}$$

We can therefore decompose our function into eigenfunctions of momentum, belonging to the eigenvalues

$$\boldsymbol{p} = \hbar \boldsymbol{k} = \left(\frac{\hbar}{L} \ell_x, \frac{\hbar}{L} \ell_y, \frac{\hbar}{L} \ell_z \right), \tag{4.73}$$

where the ℓ_α ($\alpha \in \{x, y, z\}$) are integers. The microstate of the system is therefore defined by the collection of the occupation numbers $n(\boldsymbol{p}, \sigma)$ of the single-particle states identified by the value $\boldsymbol{p}$ of the momentum and σ of the spin, that is, by the value $\hbar\sigma$ of the projection on the z axis of the intrinsic angular momentum of the particle.

It is known from quantum mechanics that there are two kinds of particles:

- Particles that have integer spin can have any occupation number. These particles are called **bosons**: photons and phonons are of this type. (More generally, elementary particles with boson-like characteristics are interaction quanta—in addition to the photon, there are particles like the W or the Z_0 that are carriers of the electroweak interaction, the gluons that are part of the strong interaction, and so on.)
- For particles with semi-integer spin, the wave function is antisymmetric with respect to the exchange of the coordinate and spin state of any two particles. Therefore, it vanishes whenever two or more particles are in the same single-particle state. This means that the occupation numbers $n(\boldsymbol{p}, \sigma)$ can assume only the values 0 and 1. This is known as the **Pauli exclusion principle**. Particles of this kind are called **fermions**. (The elementary particles with fermion-like characteristics are the basic building-blocks of matter, like protons, electrons, neutrons, and so on. The "exceptions" are the neutrinos, because of the weakness of their interactions.)

Given the complete collection of occupation numbers $\nu = (n(\boldsymbol{p}, \sigma))$ of the various single-particle states, the total energy is given by

$$\mathcal{H}(\nu) = \sum_{\boldsymbol{p},\sigma} \epsilon(\boldsymbol{p}, \sigma) n(\boldsymbol{p}, \sigma), \tag{4.74}$$

where $\epsilon(\boldsymbol{p}, \sigma)$ is the energy of a particle with momentum $\boldsymbol{p}$, and z-component of spin equal to σ, equal to $p^2/2m$ for a nonrelativistic particle. The total number of particles is obviously given by

$$N(\nu) = \sum_{\boldsymbol{p},\sigma} n(\boldsymbol{p}, \sigma). \tag{4.75}$$

Then the canonical partition function is given by

$$Z_{\mathrm{C}}(N) = \sum_{\nu} \delta\left(N - \sum_{\boldsymbol{p},\sigma} n(\boldsymbol{p}, \sigma)\right) \exp\left[-\frac{1}{k_{\mathrm{B}} T} \sum_{\boldsymbol{p},\sigma} \epsilon(\boldsymbol{p}) n(\boldsymbol{p}, \sigma)\right]. \tag{4.76}$$

The difficulty in evaluating this sum lies in the constraint of a fixed value for the total number of particles. This constraint is expressed by the delta function, which in reality is a Kronecker delta, since the number of particles is a discrete quantity. This constraint makes it impossible to calculate the sum as it stands. We are therefore forced to soften it, passing to the grand canonical ensemble. We thus evaluate the grand canonical partition function:

$$Z_{\mathrm{GC}} = \sum_{N=0}^{\infty} \exp\left(\frac{\mu N}{k_{\mathrm{B}} T}\right) Z_{\mathrm{C}}(N) = \sum_{\nu} \exp\left(-\sum_{\boldsymbol{p},\sigma} n(\boldsymbol{p}, \sigma) \frac{\epsilon(\boldsymbol{p}) - \mu}{k_{\mathrm{B}} T}\right). \tag{4.77}$$

The partition function now factorizes into sums involving the occupation numbers of the individual single-particle states:

$$Z_{\mathrm{GC}} = \prod_{\boldsymbol{p},\sigma} \sum_{n(\boldsymbol{p},\sigma)} \exp\left(-n(\boldsymbol{p},\sigma)\frac{\epsilon(\boldsymbol{p})-\mu}{k_{\mathrm{B}}T}\right) = \prod_{\boldsymbol{p},\sigma} \zeta\left(\frac{\epsilon(\boldsymbol{p},\sigma)-\mu}{k_{\mathrm{B}}T}\right), \tag{4.78}$$

where we have introduced the single-particle partition function

$$\zeta(x) = \sum_{n} \mathrm{e}^{-nx}, \tag{4.79}$$

where $n \in \{0, 1\}$ for fermions and $n \in \{0, 1, 2, \ldots\}$ for bosons.

We can now evaluate the single-particle partition function for the two kinds of particles:

- For bosons, the sum runs over all nonnegative integers, and we obtain, for $x > 0$,

$$\zeta_{-}(x) = \frac{1}{1-\mathrm{e}^{-x}}. \tag{4.80}$$

  For $x \leq 0$, the series does not converge.

- For fermions, we have only two terms, and we obtain

$$\zeta_{+}(x) = 1 + \mathrm{e}^{-x}. \tag{4.81}$$

We thus obtain the general expression of the logarithm of the grand canonical partition function:

$$\begin{aligned} \ln Z_{\mathrm{GC}} &= \sum_{\boldsymbol{p},\sigma} \ln \zeta_{\pm}\left(\frac{\epsilon(\boldsymbol{p},\sigma)-\mu}{k_{\mathrm{B}}T}\right) \\ &= (\pm)\sum_{\boldsymbol{p},\sigma} \ln\left\{1 \pm \exp\left[-\frac{\epsilon(\boldsymbol{p},\sigma)-\mu}{k_{\mathrm{B}}T}\right]\right\}. \end{aligned} \tag{4.82}$$

The upper signs apply for fermions and the lower ones for bosons.

The photon (or phonon) gas that we examined in the preceding section corresponds to the case in which the chemical potential vanishes. This is because the total number of particles is not fixed. In effect, the chemical potential is given by the derivative of the free energy (either Helmholtz's or Gibbs's) with respect to the number of particles. If this number is not fixed, thermodynamic equilibrium is reached for that value of N where these potentials reach their minimum. At equilibrium, we therefore have

$$\mu = \left.\frac{\partial F}{\partial N}\right)_{V} = 0, \tag{4.83}$$

or the analogous expression that contains G.

The spin variable σ takes $(2s+1)$ different values: $\sigma = -s, -s+1, \ldots, s-1, s$, where $\hbar^2 s(s+1)$ is the square magnitude of the intrinsic angular momentum of the particle. If $\epsilon(\boldsymbol{p},\sigma)$ does not depend on σ, the sum over the different spin states can be performed

immediately, and we are therefore left with

$$\ln Z_{\mathrm{GC}} = (\pm)(2s+1)\sum_{\boldsymbol{p}} \ln\left\{1 \pm \exp\left[-\frac{\epsilon(\boldsymbol{p},\sigma)-\mu}{k_{\mathrm{B}}T}\right]\right\}. \tag{4.84}$$

From now on, we suppose that $\epsilon(\boldsymbol{p},\sigma) = p^2/2m$. In order to calculate $\ln Z$, let us assume that no term in the last expression gives a very large contribution. This is always the case with fermions, while it is true only if μ is strictly less than zero in the case of bosons—if at the end of our calculations this condition is violated, we will have to reconsider this point. Having established this hypothesis, let us transform the sum into an integral, keeping in mind that in the volume element $\mathrm{d}\boldsymbol{p}$, there are $(L/2\pi\hbar)^3\mathrm{d}\boldsymbol{p} = (V/h^3)\mathrm{d}\boldsymbol{p}$ terms in the sum:

$$\begin{aligned}\ln Z_{\mathrm{GC}} &\simeq \frac{V}{h^3}(2s+1)\int \mathrm{d}\boldsymbol{p}\, \ln\zeta_{\pm}\left(\frac{p^2/2m-\mu}{k_{\mathrm{B}}T}\right)\\ &= (\pm)\frac{V}{h^3}(2s+1)\int_0^\infty 4\pi p^2\,\mathrm{d}p\, \ln\left[1 \pm z\exp\left(-\frac{p^2}{2mk_{\mathrm{B}}T}\right)\right],\end{aligned} \tag{4.85}$$

where, in the last step, we introduced the fugacity $z = \exp(\mu/k_{\mathrm{B}}T)$.

By performing the change of variables $p \longrightarrow x = (p^2/2mk_{\mathrm{B}}T)^{1/2}$, we obtain

$$\ln Z_{\mathrm{GC}} = (\pm)\frac{V}{h^3}(2s+1)4\pi(2mk_{\mathrm{B}}T)^{3/2}\int_0^\infty x^2\mathrm{d}x\, \ln(1 \pm z\mathrm{e}^{-x^2}). \tag{4.86}$$

The integral in x can be evaluated as follows:

$$\begin{aligned}\mathcal{I} &= (\pm)\int_0^\infty \mathrm{d}x\, x^2\, \ln\left[1 \pm z\mathrm{e}^{-x^2}\right]\\ &= (\mp)\int_0^\infty \mathrm{d}x\, x^2 \sum_{n=1}^\infty \frac{(\mp z)^n \mathrm{e}^{-nx^2}}{n}\\ &= (\mp)\sum_{n=1}^\infty \frac{(\mp z)^n}{n^{5/2}}\int_0^\infty \mathrm{d}x\, x^2\mathrm{e}^{-x^2} = \frac{\sqrt{\pi}}{4}(\mp)\phi_{5/2}(\mp z),\end{aligned} \tag{4.87}$$

where we have introduced the **polylog function**, defined by

$$\phi_\alpha(z) = \sum_{n=1}^\infty \frac{z^n}{n^\alpha}. \tag{4.88}$$

We thus obtain

$$\ln Z_{\mathrm{GC}} = \frac{V}{\lambda_{\mathrm{B}}^3}(2s+1)(\mp)\phi_{5/2}(\mp z), \tag{4.89}$$

where $\lambda_{\mathrm{B}} = (h^2/2\pi mk_{\mathrm{B}}T)^{1/2}$ is the thermal de Broglie wavelength.

The average number of particles can be obtained by taking the derivative of $\ln Z$ with respect to $\mu/k_{\mathrm{B}}T$. We obtain

$$\begin{aligned}N_{\mathrm{eq}} &= z\frac{\partial \ln Z_{\mathrm{GC}}}{\partial z} = \frac{V}{h^3}(2s+1)\int_0^\infty 4\pi p^2\mathrm{d}p\, \frac{1}{\mathrm{e}^{p^2/2mk_{\mathrm{B}}T}/z \pm 1}\\ &= \frac{V}{\lambda_{\mathrm{B}}^3}(2s+1)(\mp)\phi_{3/2}(\mp z),\end{aligned} \tag{4.90}$$

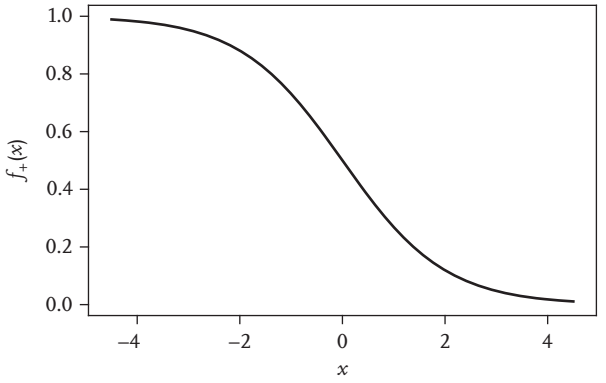


FIGURE 4.6. Fermi factor $f_+(x) = (e^x + 1)^{-1}$.

where we have exploited the relation

$$z\frac{\partial}{\partial z}\phi_\alpha(z) = z\frac{\partial}{\partial z}\sum_m \frac{z^n}{n^\alpha} = \sum_n \frac{z^n}{n^{\alpha-1}} = \phi_{\alpha-1}(z). \tag{4.91}$$

This equation can be interpreted by saying that the average occupation number of the single-particle state defined by $(\boldsymbol{p}, \sigma)$ is given for the fermions by the **Fermi factor**:

$$\langle n(\boldsymbol{p}, \sigma)\rangle = f_+\left(\frac{\epsilon(\boldsymbol{p}, \sigma) - \mu}{k_B T}\right), \tag{4.92}$$

where

$$f_+(x) = \frac{1}{e^x + 1}. \tag{4.93}$$

For the bosons we have

$$\langle n(\boldsymbol{p}, \sigma)\rangle = f_-\left(\frac{\epsilon(\boldsymbol{p}, \sigma) - \mu}{k_B T}\right), \tag{4.94}$$

where $f_-(x) = (e^x - 1)^{-1}$ is the Bose factor that we have already encountered. The Bose factor is shown in fig. 4.2, while the Fermi factor is shown in fig. 4.6. It approaches 1 for $x \to -\infty$, decreases monotonically with x, assumes the value $1/2$ for $x = 0$, and behaves as e^{-x} for $x \gg 1$.

In order to obtain the equation of state, it is necessary to solve eq. (4.90) with respect to z (considering N as known) and substitute in eq. (4.89), since, as we have seen in sec. 3.16, we have for a simple fluid

$$\ln Z_{GC} = \frac{\Omega}{k_B T} = \frac{pV}{k_B T}. \tag{4.95}$$

In practice, we can make a parametric plot of $p\lambda_B^3/k_B T$ from eq. (4.89) vs. $\rho\lambda_B^3$ from eq. (4.90) as a function of z. Notice that, although the series that defines the polylog function does not converge for $z > 1$, the integrals that appear in both these equations are well

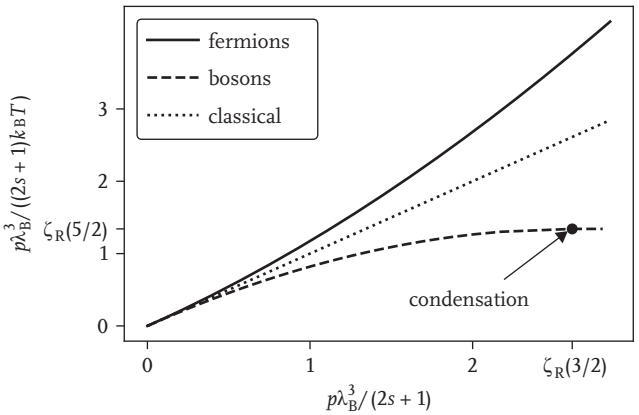


FIGURE 4.7. Equations of state of the ideal quantum gases. The graph represents the quantity $p\lambda_B^3/((2s+1)k_BT)$, where p is the pressure and λ_B the thermal de Broglie wavelength, plotted versus $\rho\lambda_B^3/(2s+1)$, where $\rho = N/V$ is the particle density. The classical equation of state is also shown. The condensation point is discussed in sec. 4.9.

behaved for the fermion case. The boson case requires more attention, as we discuss in sec. 4.9. We thus obtain the equations of state shown in fig. 4.7. We notice that the pressure is always larger for fermions and smaller for bosons than the corresponding classical one, for the same value of the particle density (measured in units λ_B^{-3}).

The classical limit of these relations can be obtained by considering small values of z. In this case, if we stop at the first term of the expansion of $\phi_{3/2}(\mp z) \simeq \mp z$, we obtain, for both bosons and fermions,

$$z \simeq \frac{N\lambda_B^3}{(2s+1)V}. \tag{4.96}$$

Substituting this result in the expression for pressure, and since $\phi_{5/2}(z) \simeq z$ to first order, we once again obtain the ideal gas law in the form $p/k_BT = \rho$. We can therefore expect the classical theory to be reliable when the number ξ of particles contained in a volume equal to the cube of the thermal de Broglie wavelength is small. This quantity is given by

$$\xi = \frac{N\lambda_B^3}{V}. \tag{4.97}$$

The classical results are reliable only for $\xi \ll 1$. Otherwise, the quantum properties become evident, and the gas is said to be **degenerate**. It is easy (and we leave it as an exercise) to calculate the first quantum corrections to the equation of ideal gases—one should keep in mind that usually the corrections due to interactions are much more important than those of quantum origin.

Exercise 4.4 Evaluate the corrections to the classical equation of state of the quantum gases to order ξ^2, where ξ is the degeneration parameter defined in eq. (4.97).

4.4 Electrons in Metals

Electrons in metals form an approximately ideal and very degenerate fermion gas—in other words, one whose behavior is very far from classical. The first statement is surprising, since electrons are subject to strong Coulomb interactions with the atomic nuclei of the lattice, as well as with other electrons. It is beyond the scope of this book to explain why, as a first approximation, one can treat the electrons as being interaction-free, although with a non-trivial dependence between their momentum and their energy. Insofar as interactions with the atoms of the lattice are concerned, since they are represented by a spatially periodic potential, their effect can be summarized by changing the expression of the electron energy to a function of the wave vector. We here refer the reader to any good book on solid state—for example, Ashcroft and Mermin [Ashc76]—to understand how the effect of the periodic potential is to break up the intervals of the allowable values for single-particle energy into bands, and how each band is run over when the particle's wave vector varies in the region known as the **first Brillouin zone**.

It is easy, on the other hand, to calculate the characteristic values of the degeneration parameter $\xi = N\lambda_B^3/V$. In the case of electrons, we have $m = 9.11 \cdot 10^{-31}$ kg, and therefore, at room temperature (300 K), the thermal de Broglie wavelength is given by

$$\begin{aligned}\lambda_B &= \left(\frac{h^2}{2\pi m k_B T}\right)^{1/2} \simeq \left[\frac{(6.63 \cdot 10^{-34})^2}{2 \cdot 3.14 \cdot 9.11 \cdot 10^{-31} \cdot 1.38 \cdot 10^{-23} \cdot 300}\right]^{1/2} \\ &= 4.31 \cdot 10^{-9}\,\mathrm{m} = 43.1\,\text{Å}.\end{aligned} \tag{4.98}$$

As a result, if we count roughly one free electron per atom, and since atomic distances in solids are of the order of Ångströms, the value of ξ will be of the order of $40^3 \simeq 6.4 \cdot 10^4$, and the gas will therefore be very far from the classical limit.

Let us therefore attempt to investigate the behavior of a strongly degenerate fermion gas. In order to clarify our ideas, let us assume that the particles are free so that the relation between the particle momentum $\boldsymbol{p}$ and its energy ϵ is the usual one:

$$\epsilon(\boldsymbol{p}) = \frac{p^2}{2m}. \tag{4.99}$$

In the limit $T \to 0$, the Fermi factor tends to 1 if $\epsilon - \mu < 0$, and it tends to 0 if $\epsilon - \mu > 0$—this means that all available states with energy below the chemical potential are completely occupied, while those at higher energies are empty. The maximum value of the energy of the occupied states at $T = 0$ is called the **Fermi energy** and is denoted by ϵ_F.

Let us define the **density of states** $\omega(\epsilon)$ such that $\omega(\epsilon)\,\mathrm{d}\epsilon$ is the number per unit volume of single-particle states with a given spin, with energy situated between ϵ and $\epsilon + \mathrm{d}\epsilon$. It is given by

$$\omega(\epsilon) = \frac{1}{(2s+1)V} \sum_{\boldsymbol{p},\sigma} \delta\left(\epsilon - \epsilon(\boldsymbol{p},\sigma)\right). \tag{4.100}$$

In the case of free electrons, one gets

$$\omega(\epsilon) = \frac{4\pi m (2m\epsilon)^{1/2}}{h^3}. \tag{4.101}$$

The total number of particles at $T = 0$ is expressed as a function of the Fermi energy ϵ_F by

$$N = V(2s+1)\int_0^{\epsilon_F} d\epsilon\, \omega(\epsilon). \tag{4.102}$$

For free electrons, we obtain

$$N = \frac{4\pi V}{3h^3}(2s+1)\,(2m\epsilon_F)^{3/2}. \tag{4.103}$$

By inverting this relation, we obtain the expression for the Fermi energy:

$$\epsilon_F = \frac{h^2}{2m}\left[\frac{3}{4\pi}\cdot\frac{N}{V(2s+1)}\right]^{2/3}. \tag{4.104}$$

It is interesting to note that the Fermi energy is inversely proportional to the fermion mass—and this is the reason why phenomena related to the Fermi distribution are more obvious in electrons. If the density of states at the Fermi energy does not vanish, as for the case of free electrons, the chemical potential is equal to the Fermi energy at $T = 0$.

We can associate a temperature T_F, which is called the **Fermi temperature**, to the Fermi energy by means of the relation

$$\epsilon_F = k_B T_F. \tag{4.105}$$

Then, another way to express the degree of quantum degeneration of electrons in metals is to consider the ratio T_F/T. The values characteristic of T_F for metals go from the $1.84 \cdot 10^4$ K of cesium to the $16.0 \cdot 10^4$ K of beryllium.

In order to calculate the properties of fermion gas at low temperatures, one uses a technique developed by Sommerfeld. From now on, we will make use of the fact that, for electrons, one has $s = 1/2$, and therefore $2s + 1 = 2$. Define the integral of the density of states $\Omega(\omega)$ via the expression:

$$\Omega(\epsilon) = \int_0^{\epsilon} d\epsilon'\, \omega(\epsilon'). \tag{4.106}$$

We can then evaluate the average number of particles as a function of the chemical potential μ for small values of T as follows:

$$\begin{aligned} N &= 2V\int_0^{\infty} d\epsilon\, \frac{d\Omega}{d\epsilon} f_+\left(\frac{\epsilon-\mu}{k_B T}\right) \\ &= 2V\left[\Omega(\epsilon) f_+\left(\frac{\epsilon-\mu}{k_B T}\right)\right]_0^{\infty} - \frac{2V}{k_B T}\int_0^{\infty} d\epsilon\, \Omega(\epsilon) f_+'\left(\frac{\epsilon-\mu}{k_B T}\right). \end{aligned} \tag{4.107}$$

The first term vanishes, while in the second term, the $\Omega(\epsilon)$ can be expanded into powers of $(\epsilon - \mu)$, since the derivative of the Fermi factor differs from zero only within a small interval on the order of $k_B T$ around the origin. We thus obtain

$$N = -\frac{2V}{k_B T}\int_0^{\infty} d\epsilon \sum_{n=0}^{\infty} \frac{1}{n!}\left.\frac{d^n\Omega}{d\epsilon^n}\right|_{\epsilon=\mu} (\epsilon-\mu)^n f_+'\left(\frac{\epsilon-\mu}{k_B T}\right). \tag{4.108}$$

We can therefore extend the integration over ϵ to the entire real axis. Since $f'_+(x)$ is an even function of x, only terms with even n survive in this expression, and yield

$$\begin{aligned} N &= -\frac{2V}{k_B T} \sum_{n=0}^{\infty} \frac{1}{(2n)!} \left. \frac{\mathrm{d}^{2n}\Omega}{\mathrm{d}\epsilon^{2n}} \right|_{\epsilon=\mu} \int_{-\infty}^{+\infty} \mathrm{d}\epsilon \, (\epsilon-\mu)^{2n} f'_+ \left(\frac{\epsilon-\mu}{k_B T} \right) \\ &= 2V\Omega(\mu) + 2V \sum_{n=1}^{\infty} \frac{(k_B T)^{2n}}{(2n)!} \left. \frac{\mathrm{d}^{2n}\Omega}{\mathrm{d}\epsilon^{2n}} \right|_{\epsilon=\mu} \mathcal{I}_n, \end{aligned} \tag{4.109}$$

where we have defined

$$\mathcal{I}_n = \int_{-\infty}^{+\infty} \mathrm{d}x \, \frac{x^{2n} \, \mathrm{e}^x}{(\mathrm{e}^x+1)^2}. \tag{4.110}$$

These integrals $\mathcal{I}_n$ can be evaluated explicitly, but their values do not interest us here, except for $n=1$, for which we have $\mathcal{I}_1 = \zeta_R(2) = \pi^2/6$. To obtain the chemical potential, we keep the first nontrivial order in eq. (4.109) and solve for μ at the given value of N. We thus obtain

$$\mu = \epsilon_F - \frac{(k_B T)^2}{2} \zeta_R(2) \frac{\omega'(\epsilon_F)}{\omega(\epsilon_F)}. \tag{4.111}$$

To evaluate the specific heat, we take advantage of the expression (4.92) of the average occupation number of a single particle state. We can then express the free energy in terms of the Fermi function and of the density of states:

$$E = 2V \int_0^{\infty} \mathrm{d}\epsilon \, \epsilon \, \omega(\epsilon) f_+ \left(\frac{\epsilon-\mu}{k_B T} \right). \tag{4.112}$$

Let us define the function

$$\Omega_E(\epsilon) = \int_0^{\epsilon} \mathrm{d}\epsilon' \, \epsilon' \omega(\epsilon'). \tag{4.113}$$

We can then apply the same method as above to obtain

$$\begin{aligned} E &= 2V\Omega_E(\mu) + 2V \sum_{n=1}^{\infty} (k_B T)^{2n} \mathcal{I}_n \left. \frac{\mathrm{d}^{2n}\Omega_E}{\mathrm{d}\epsilon^{2n}} \right|_{\epsilon=\mu} \\ &\simeq 2V \left[\int_0^{\mu} \mathrm{d}\epsilon \, \epsilon \, \omega(\epsilon) + \frac{(k_B T)^2}{2} \mathcal{I}_1 \left. \frac{\mathrm{d}}{\mathrm{d}\epsilon} \left(\epsilon \, \omega(\epsilon) \right) \right|_{\epsilon=\mu} \right] \\ &= 2V \left[\int_0^{\mu} \mathrm{d}\epsilon \, \epsilon \, \omega(\epsilon) + \frac{(k_B T)^2}{2} \mathcal{I}_1 \left(\omega(\mu) + \mu \, \omega'(\mu) \right) \right] \\ &\simeq 2V \left[\int_0^{\epsilon_F} \mathrm{d}\epsilon \, \epsilon \, \omega(\epsilon) + \frac{(k_B T)^2}{2} \frac{\pi^2}{6} \omega(\epsilon_F) \right], \end{aligned} \tag{4.114}$$

where we have taken into account the shift in the chemical potential given by eq. (4.111). One can interpret intuitively this result by considering that at a small temperature T only those electrons whose energy lies below ϵ_F by $k_B T$ will be thermally excited. This yields a number of electrons proportional to $\omega(\epsilon_F) \, k_B T$ per unit volume, which receive each an

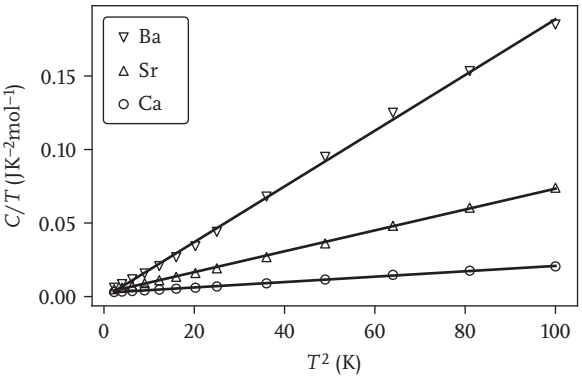


FIGURE 4.8. Plot of C/T vs. T^2 for Ba, Sr, and Ca. Lines are unsophisticated fits to the data. Data from [Robe57].

amount of energy of order $k_B T$. Therefore we expect that $E(T) - E(0) \propto (k_B T)^2$, where the proportionality constant is of order $V\omega(\epsilon_F)$.

Taking the derivative, we obtain that the specific heat is linear in T and is given by

$$C = \frac{\pi^2 V}{3} k_B^2 T\, \omega(\epsilon_F). \tag{4.115}$$

This expression holds whatever the form of $\epsilon(\boldsymbol{p}, \sigma)$.

In the case of free electrons, by using eqs. (4.101) and (4.104), one obtains

$$C = \frac{\pi^2}{2} \frac{T}{T_F} N k_B. \tag{4.116}$$

Since this contribution is inversely proportional to the Fermi temperature, which is usually quite large for metals, the contribution of conduction electrons to the specific heat is usually negligible at room temperature.

Since the specific heat of a metal at low temperatures is given by the sum of the electronic and lattice contributions, for small values of T it behaves as

$$C = aT + bT^3, \tag{4.117}$$

where a and b are respectively related to the density of states at the Fermi surface and to the Debye temperature. Thus the electronic contribution dominates at low temperatures. It is possible to evaluate both $\omega(\epsilon_F)$ and T_D by plotting C/T vs. T^2. Such a plot is shown in fig. 4.8, with the data of [Robe57]. Although the specific heat of the electron gas also exhibits a term proportional to T^3, this contribution is inversely proportional to a power of the Fermi temperature, and is therefore negligible with respect to the lattice contribution for all practical purposes.

Exercise 4.5 (Pauli paramagnetism) The paramagnetic properties of a number of metals can be explained by considering the effect of a magnetic field on a degenerate electron gas.

Assume that each electron has a magnetic dipole of magnitude μ_{m} parallel to its spin, and that a magnetic field of induction vector $\boldsymbol{B}$ is applied to the system. Then states with spin parallel or antiparallel to the field have different energies.

1. Evaluate the total magnetization $\boldsymbol{M}$ of the system as a function of $\boldsymbol{H}$ at zero temperature for a three-dimensional system of N free electrons.
2. Same question for a two-dimensional system.

4.5 Relation between Pressure and Internal Energy

In all the cases in which the relation between wave vector (or momentum) and energy in single-particle states takes the form of a power law, it is possible to prove a proportionality relation between the pressure and the internal energy density in ideal gases (classical or quantum). For generality's sake, we will consider particles in d-dimensional space that possess a relation between energy and momentum of the form

$$\epsilon(\boldsymbol{p},\sigma) = \kappa \left|\boldsymbol{p}\right|^{\alpha}, \tag{4.118}$$

where κ is a constant. In the case of particles with mass, such as electrons, one has $\alpha = 2$ and $\kappa = 1/(2m)$; in the case of photons (or of ultrarelativistic particles), one has $\alpha = 1$ and $\kappa = c$ (where c is the speed of light). A similar relation with c_{s}, that is, the speed of sound, holds for phonons.

We have seen in sec. 3.16 that the pressure is given by the logarithm of the grand canonical partition function, which for our case is expressed by

$$\frac{pV}{k_{\mathrm{B}}T} = \ln Z = (\pm)\frac{V}{h^3}(2s+1)\int_0^{\infty} S_d p^{d-1}\,\mathrm{d}p\,\ln\left[1 \pm z\exp\left(-\frac{\epsilon(\boldsymbol{p})}{k_{\mathrm{B}}T}\right)\right], \tag{4.119}$$

where $z = \mathrm{e}^{\mu/k_{\mathrm{B}}T}$ is the fugacity and S_d is the area of the unit sphere in d dimensions. By performing the change of variables

$$p \longrightarrow x = \left(\frac{\kappa p^{\alpha}}{k_{\mathrm{B}}T}\right), \tag{4.120}$$

we obtain

$$\ln Z = \frac{V}{h^d}\left(\frac{k_{\mathrm{B}}T}{\kappa}\right)^{d/\alpha}\frac{S_d}{\alpha}\int_0^{\infty}\mathrm{d}x\,x^{(d/\alpha)-1}(\pm)\ln(1 \pm z\mathrm{e}^{-x}). \tag{4.121}$$

On the other hand, the internal energy is given by

$$E = -\left.\frac{\partial \ln Z}{\partial(1/k_{\mathrm{B}}T)}\right)_{V,z}. \tag{4.122}$$

Since $\ln Z$, when V and z are kept constant, is proportional to $(1/k_{\mathrm{B}}T)^{-d/\alpha}$, we obtain

$$E = -\left.\frac{\partial \ln Z}{\partial(1/k_{\mathrm{B}}T)}\right)_{V,z} = -\left(-\frac{d}{\alpha}\right)\left(\frac{1}{k_{\mathrm{B}}T}\right)^{-1}\ln Z = \frac{d}{\alpha}k_{\mathrm{B}}T\ln Z = \frac{d}{\alpha}pV, \tag{4.123}$$

which is the result we wanted to prove. This result is independent of the nature (fermion or boson) of the particles being considered.

More particularly, in the case of electrons at high densities, the internal energy is of the order of the Fermi energy, which, as we have seen, is proportional to the density to 2/3. We thus obtain

$$pV \propto V^{2/3}, \tag{4.124}$$

corresponding to

$$pV^{5/3} = \text{const.} \tag{4.125}$$

This equation recalls the law of adiabatics of the classical gas, $pV^{\gamma} = \text{const.}$ where $\gamma = C_p/C_V$, eq. (2.160). Indeed, Nernst's postulate implies that adiabatics and isotherms coincide at $T = 0$, and one can show that for our system one has $C_p/C_V = 1 + d/\alpha$. The derivation of this relation is left to the reader as an exercise.

Exercise 4.6 Show that for a quantum gas in d dimensions, whose single-particle energy $\epsilon(\boldsymbol{p})$ is proportional to p^{α}, one has $C_p/C_V = 1 + d/\alpha$.

Exercise 4.7 (Fermions in two dimensions) Let us consider a system of noninteracting fermions of spin 1/2 and mass m in two dimensions, contained in a square box of side L.

1. Evaluate the density of states $\omega(\epsilon) = \sum_{k,\sigma} \delta(\epsilon - \epsilon(\boldsymbol{k}, \sigma))$, where $\boldsymbol{k}$ are the allowed values of the wave vector, σ the possible values of the spin, and $\epsilon(\boldsymbol{k}, \sigma)$ the corresponding single-particle energy.
2. Evaluate the Fermi energy ϵ_{F} as a function of the particle density $\rho = N/V$.
3. Evaluate the equation of state $p(\rho)$ at vanishing temperature.
4. Show that the chemical potential μ tends exponentially toward ϵ_{F} as $T \to 0$.

Exercise 4.8 (Neutrino density in the universe) Current cosmological theories predict not only a uniform background radiation density in thermal equilibrium at the temperature of 2.725 K, but also an equilibrium distribution of neutrinos at the same temperature.

Neutrinos are spin–$\frac{1}{2}$ particles (and therefore fermions), but one finds experimentally that their spin is always *parallel* to their momentum $\boldsymbol{p}$. Thus there is only one possible state for the spin for each value of the momentum. Moreover, their mass is indistinguishable from zero and therefore the energy of a state of momentum $\boldsymbol{p}$ is equal to $c\,|\boldsymbol{p}|$, where c is the speed of light. Their (weak) interaction with matter implies that the number of neutrinos is not fixed.

Evaluate:

1. The number of neutrinos per unit volume;
2. their average energy;
3. their energy spectrum.

Data:

- $c \simeq 3.00 \cdot 10^8\,\mathrm{ms}^{-1}$;
- $k_\mathrm{B} \simeq 1.38 \cdot 10^{-23}\,\mathrm{JK}^{-1}$;
- $\int_0^\infty x^2 \mathrm{d}x/(1+\mathrm{e}^x) = \frac{3}{2}\zeta_\mathrm{R}(3) \simeq 1.80$;
- $\int_0^\infty x^3 \mathrm{d}x/(1+\mathrm{e}^x) = 7\pi^4/120 \simeq 5.67$.

4.6 Diamagnetism

Most materials that do not possess unmatched electrons exhibit diamagnetism, in that their magnetic moments tend to align opposite to external applied fields. This property has been highlighted in 1997, when a group of researchers in the Netherlands and England made a frog levitate in a magnetic field [Berr97]. The feat earned the team the 2000 physics Ig Nobel Prize (one of the authors later got the Nobel prize for a different work). Indeed, a diamagnetic material is pushed in the direction in which the field's magnitude decreases. For systems mainly composed of water (like frogs) and devices of a fraction of a meter in size, this requires very intense fields, of the order of 16 T. On the other hand, the Bohr-van Leeuwen theorem (see exercise 3.3.2) prevents diamagnetism from appearing at a classical level. In order to understand the origins of diamagnetism it is necessary to resort to quantum mechanics.

4.6.1 Langevin Diamagnetism

In materials composed of atoms with complete electron layers, one can understand diamagnetism by looking at the effect of the applied field on independent atoms. Let us then consider N such atoms placed, say, on a crystal lattice, and subject to a magnetic field $\boldsymbol{B}$ directed along the z axis. Then the Hamiltonian of a single atom is given by

$$\hat{\mathcal{H}} = \sum_i \left[\frac{1}{2m} \left(\boldsymbol{p} + e\boldsymbol{A}(\boldsymbol{r}_i)\right)^2 + 2\mu_\mathrm{B} \boldsymbol{B} \cdot \boldsymbol{S} \right] + U(\boldsymbol{r}), \tag{4.126}$$

where the sum runs over all electrons in the atom, $\boldsymbol{p} = (\boldsymbol{p}_i)$ are the momenta and $\boldsymbol{r} = (\boldsymbol{r}_i)$ are the position vectors of the electrons, $\boldsymbol{S} = (\boldsymbol{S}_i)$ are their spins, m is the electron mass, $-e$ is its charge, $\mu_\mathrm{B} = e\hbar/(2m)$ is Bohr's magneton, and $U(\boldsymbol{r})$ is the electrostatic energy of the system. Here $\boldsymbol{A}(\boldsymbol{r})$ is the vector potential related to $\boldsymbol{B}$ by $\boldsymbol{B} = \nabla \times \boldsymbol{A}$, which for a uniform

field parallel to the z axis and of magnitude B can be chosen as follows:

$$A_x = -\frac{1}{2} yB, \qquad A_y = \frac{1}{2} xB, \qquad A_z = 0. \tag{4.127}$$

Substituting this expression in eq. (4.126) and remembering that $\boldsymbol{p} = -\mathrm{i}\hbar\nabla$, we obtain

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 + \mu_B \left(L_z + 2S_z\right) B + \frac{e^2}{8m} B^2 \sum_i \left(x_i^2 + y_i^2\right). \tag{4.128}$$

Here L_z is the z component of the total kinetic moment of the atom, S_z the z component of its total spin, and $\mathcal{H}_0$ the Hamiltonian of the atom in the absence of the applied field.

The intensity of the usually applied magnetic fields (even at 16 T) is usually too small to be able to excite the atom above its electronic ground state. We shall thus suppose that the atom remains in its ground state even if $B > 0$. In this state, one has $|\boldsymbol{L}| = |\boldsymbol{S}| = 0$, and only the quadratic term in B of eq. (4.128) survives. To lowest order in the perturbation, the ground-state energy is therefore given by

$$E_0(B) = E_0(0) + \frac{e^2 B^2}{8m} \left\langle 0 \left| \sum_i (x_i^2 + y_i^2) \right| 0 \right\rangle + \cdots, \tag{4.129}$$

where $|0\rangle$ denotes the ground state in the absence of the applied field.

In its fundamental state, the electron wavefunction is spherically symmetric: one thus has

$$\left\langle 0 \left| \sum_i (x_i^2 + y_i^2) \right| 0 \right\rangle = \frac{2}{3} \left\langle 0 \left| \sum_i (x_i^2 + y_i^2 + z_i^2) \right| 0 \right\rangle = \frac{2}{3} \left\langle 0 \left| \sum_i r_i^2 \right| 0 \right\rangle = \frac{2}{3} Z \left\langle r^2 \right\rangle, \tag{4.130}$$

where Z is the number of electrons (that is, the atomic number of the atoms), and $\langle r^2 \rangle$ is the average of r^2 evaluated over the electron density of the atom. We have therefore

$$E_0(B) \simeq E_0(0) + \frac{e^2 Z B^2}{12m} \left\langle r^2 \right\rangle. \tag{4.131}$$

If the temperature and the magnetic field are not too large, the atom remains in its ground state: thus its entropy remains equal to 0, and therefore its Helmholtz free energy remains equal to its internal energy, in our case $E_0(B)$. Therefore the free energy of our system of N particles is given by

$$F(T, B) = N \left[E_0 + \frac{e^2 Z B^2}{12m} \left\langle r^2 \right\rangle \right], \tag{4.132}$$

independent of the temperature. Thus we can straightforwardly obtain the magnetization

$$M = -\frac{\partial F}{\partial B} = -\left(\frac{e^2 N Z}{6m} \left\langle r^2 \right\rangle \right) B \tag{4.133}$$

and the susceptibility per particle

$$\chi = \frac{1}{N} \frac{\partial M}{\partial B} = -\frac{e^2 Z}{6m} \left\langle r^2 \right\rangle < 0. \tag{4.134}$$

In the case of a heterogeneous material, the different atomic species contribute separately to the magnetic susceptibility. One can notice that, in this calculation, statistical mechanics played no role.

Exercise 4.9 (Classical theory of diamagnetism) Quite surprisingly, the same result was obtained by P. Langevin in 1905 via a completely classical reasoning, which we shall follow in the present exercise.

Consider a classical atom subject to a uniform magnetic field $\boldsymbol{B}$ of magnitude B directed along the z axis.

1. According to the Larmor theorem, the system undergoes a precession at the angular frequency ω_{L}. Evaluate this frequency.
2. This precession is equivalent to an added circular current. Evaluate the added current density at point $\boldsymbol{r}$ assuming the electron density $\rho(\boldsymbol{r})$ to be given and to be spherically symmetric.
3. Using this result, evaluate the total magnetic moment of the atom, again exploiting spherical symmetry, and compare this result with the preceding one.

In this derivation, where did we violate the assumptions of the Bohr-van Leeuwen theorem?

4.6.2 Landau Diamagnetism

Let us now consider, following L. D. Landau [Land30], a system of N free electrons (in a uniform positive background) submitted to a uniform field B directed along the z axis. The Hamiltonian of the system is given, with the usual conventions, by

$$\hat{\mathcal{H}} = \sum_i \left[\frac{1}{2m} \left(\boldsymbol{p}_i + e\boldsymbol{A}(\boldsymbol{r}_i) \right)^2 \right] + \sum_i [g_e \mu_{\mathrm{B}} S_{zi} B]. \tag{4.135}$$

In the second term, S_{zi} is the z component of the spin of electron i, and the factor $g_e \simeq 2$ is the gyromagnetic factor of the electron. This term commutes with the previous term, therefore its contribution to the partition function factorizes. Its effect is the Ising paramagnetism that we have discussed in sec. 3.11. We shall now dwell on the first term, and we shall see shortly that it yields a diamagnetic contribution. We choose a different gauge for $\boldsymbol{A}$ than before, namely,

$$A_x = 0, \qquad A_y = Bx, \qquad A_z = 0. \tag{4.136}$$

We thus obtain the single-particle Hamiltonian

$$\hat{\mathcal{H}}_{\mathrm{sp}} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + \frac{1}{2m}\left(-\mathrm{i}\frac{\partial}{\partial y} + eBx \right)^2. \tag{4.137}$$

Let us look for eigenfunctions of the form

$$\psi(x, y, z) = \mathrm{e}^{\mathrm{i}k_y y + \mathrm{i}k_z z} \varphi(x). \tag{4.138}$$

We obtain the Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi+\frac{eB^2}{2m}\left(x+\frac{hk_y}{eB}\right)^2\varphi=\left(\epsilon-\frac{\hbar^2k_z^2}{2m}\right)\varphi. \tag{4.139}$$

This is the equation of a harmonic oscillator on the x axis of frequency $\omega_B = eB/m$, centered at the point $x_0 = -\hbar k_y/eB$. We thus obtain

$$\epsilon-\frac{\hbar^2k_z^2}{2m}=\left(n_x+\frac{1}{2}\right)\frac{e\hbar}{m}B, \qquad n_x=0,1,2,\ldots. \tag{4.140}$$

Keeping in mind that $e\hbar/m = 2\mu_\mathrm{B}$, we obtain the energy eigenvalues:

$$\epsilon(k_z,n_x)=\frac{\hbar^2k_z^2}{2m}+(2n_x+1)\,\mu_\mathrm{B}B, \qquad n_x=0,1,2,\ldots. \tag{4.141}$$

They correspond to a free motion in the z-direction, but to quantized levels (the **Landau levels**) in the xy plane. Their energy does not depend on k_y, but the oscillator center x_0 does. Let us now suppose that the electron gas is contained in a cubic box of side L with periodic boundary conditions. Then the wavenumbers k_y, k_z must be of the form

$$k_\alpha=\frac{2\pi}{L}n_\alpha, \qquad n_\alpha=0,\pm1,\pm2,\ldots \qquad \alpha=y,z. \tag{4.142}$$

The wavefunction $\varphi(x)$ has limited range, but we must impose that it lies entirely inside the box. This requires that $0 < x_0 = \hbar\left|k_y\right|/eB < L$ (neglecting corrections close to the boundary), and therefore that each state with given values of k_z and n_x have multiplicity given by

$$\omega(B)=\frac{eBL}{\hbar}\cdot\frac{L}{2\pi}=\frac{eBL^2}{h}. \tag{4.143}$$

This degeneracy is proportional to L^2, and therefore is quite large, as soon as B has a reasonable value.

We can now evaluate the single-particle partition function:

$$\begin{aligned}\zeta&=\frac{eBL^2}{h}\sum_{n_z,n_x}\mathrm{e}^{-\epsilon(n_z,n_x)/k_\mathrm{B}T}\\&=\frac{eBL^2}{h}\sum_{n_z,n_x}\exp\left[-\left(\frac{h^2n_z^2}{2mL^2\,k_\mathrm{B}T}\right)-(2n_x+1)\,\frac{\mu_\mathrm{B}B}{k_\mathrm{B}T}\right],\end{aligned} \tag{4.144}$$

The sum over n_z can be approximated by a classical Gaussian integral, and the sum over n_x is the usual geometric sum of the harmonic oscillator. Summing up, we obtain

$$\begin{aligned}\zeta&=\frac{eBL^2}{h}\left(\frac{2\pi mk_\mathrm{B}TL^2}{h^2}\right)^{1/2}\frac{\exp(-\mu_\mathrm{B}B/k_\mathrm{B}T)}{1-\exp\left(-2\mu_\mathrm{B}B/k_\mathrm{B}T\right)}\\&=V\left(\frac{2\pi mk_\mathrm{B}T}{h^2}\right)^{3/2}\frac{\mu_\mathrm{B}B/k_\mathrm{B}T}{\sinh(\mu_\mathrm{B}B/k_\mathrm{B}T)}.\end{aligned} \tag{4.145}$$

The partition function of the N-particle system is then given by

$$Z = \frac{\zeta^N}{N!}, \tag{4.146}$$

taking into account the indistinguishability of the particles. We therefore obtain

$$F = -k_B T \ln Z = -N k_B T \left[\ln \left(\frac{eV}{N \lambda_B^3} \right) + \ln \frac{\mu_B B / k_B T}{\sinh(\mu_B B / k_B T)} \right]. \tag{4.147}$$

We thus obtain the magnetization:

$$M = -\frac{\partial F}{\partial B} = -N \mu_B \, \phi_L \left(\frac{\mu_B B_0}{k_B T} \right), \tag{4.148}$$

where $\phi_L(x)$ is the **Langevin function**, defined by

$$\phi_L(x) = \coth x - \frac{1}{x} \simeq \frac{x}{3} - \frac{x^3}{45} + \cdots . \tag{4.149}$$

The diamagnetic susceptibility per particle, for $\mu_B B \ll k_B T$, is therefore given by

$$\chi_d \simeq = -\frac{\mu_B^2}{3 k_B T}, \tag{4.150}$$

and must be compared with the paramagnetic susceptibility, due to the electron spin, obtained in sec. 3.11. Taking into account that for electrons one has $\mu_m = g_e \mu_B / 2 \simeq \mu_B$, one obtains

$$\chi_p \simeq \frac{\mu_B^2}{k_B T}. \tag{4.151}$$

The paramagnetic susceptibility is therefore larger in magnitude than the diamagnetic one, but they have the same order of magnitude.

4.6.3 De Haas-van Alphen Effect

When $k_B T \ll \mu_B B$, the system of N electrons exhibits a curious behavior, known as the **de Haas-van Alphen** effect. We shall discuss it in a simplified setting, considering an electron gas in two dimensions. In this case, the single-particle energy depends only on n_x and is given by

$$\epsilon(n_x) = (2n_x + 1) \mu_B B. \tag{4.152}$$

Its degeneracy $\omega(B)$ can be written in the form

$$\omega(B) = N \frac{B}{B^*} = N w, \tag{4.153}$$

where $w = B / B^*$ and the characteristic field B^* is given by

$$B^* = \frac{N h}{e L^2} = e h \, \rho, \tag{4.154}$$

where $\rho = N/V$ is the particle density of the electron gas. From now on, we suppose that the temperature equals zero. All Landau levels below the Fermi energy ϵ_F are occupied. When $B/B^* > 1$, $\epsilon_F > \epsilon(0)$, and thus all electrons can stay in the ground state, and the energy per particle is given by

$$\frac{E}{N} = \epsilon(0) = \mu_B B. \tag{4.155}$$

As soon as $B/B^* < 1$, some electrons will start occupying the $n_x = 1$ state. As the field decreases, the $n_x = 1$ state fills up, being complete when $B/B^* = \frac{1}{2}$. The same reasoning can be applied to higher Landau levels. Suppose that B is such that the nth level, but not the $(n+1)$th one, is filled. This happens when

$$\omega(B)(n+1) < N < \omega(B)(n+2). \tag{4.156}$$

This corresponds to

$$\frac{1}{n+2} < \frac{B}{B^*} < \frac{1}{n+1}. \tag{4.157}$$

In this range, one has

$$\begin{aligned} \frac{E}{N} &= w \sum_{k=0}^{n} \epsilon(k) + [1 - (n+1)w]\,\epsilon_{n+1} \\ &= \mu_B B\,[2n + 3 - w(n+1)(n+2)]. \end{aligned} \tag{4.158}$$

We obtain therefore

$$\frac{E(B)}{N} = \mu_B B \cdot \Phi\left(\frac{B}{B^*}\right), \tag{4.159}$$

where

$$\Phi(w) = \begin{cases} w, & \text{for } w > 1, \\ w[2n + 3 - w(n+1)(n+2)], & \text{for } n+1 < 1/w < n+2, \\ & n = 0, 1, 2, \ldots. \end{cases} \tag{4.160}$$

One can straightforwardly obtain the magnetization per particle:

$$\frac{M(B)}{N} = \mu_B \mathcal{M}\left(\frac{B}{B^*}\right), \tag{4.161}$$

where

$$\mathcal{M}(w) = \begin{cases} -1, & \text{for } w > 1, \\ 2(n+1)(n+2)w - (2n+3), & \text{for } n+1 < 1/w < n+2, \\ & n = 0, 1, 2, \ldots. \end{cases} \tag{4.162}$$

We thus witness periodic jumps when B^*/B crosses the integers, connected by linear ramps, as shown in fig. 4.9. As a consequence, the susceptibility exhibits plateaus most of the time, with value $2(n+1)(n+2)\mu_B/B^*$, excluding the delta-like singularities when B^*/B is integer.

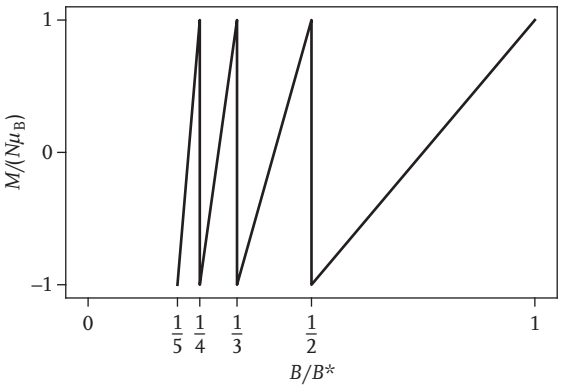


FIGURE 4.9. De Haas-van Alphen effect. Magnetization per particle M/N (in units of μ_B) vs. B/B^* at zero temperature. One notices the jumps when B^*/B crosses an integer value, and the linear ramps in between. Notice that there is no "weak-field limit," at least for $T = 0$.

The practical interest of the de Haas-van Alphen effect lies in its use for measuring the Fermi surface. Indeed, L. Onsager pointed out in 1952 [Onsa52] that the oscillation period in $1/B$ is given by the simple relation

$$\Delta\left(\frac{1}{B}\right) = \frac{2\pi e}{\hbar A_\perp}, \tag{4.163}$$

where $A_\perp$ is the area of an extremal cross section of the Fermi surface normal to the direction of the field. I refer you to the textbooks on solid state physics (such as [Merm66, ch. 14]) for a more detailed discussion.

4.7 White Dwarfs

When a sun-like star exhausts its fuel, after a number of events that will not be discussed here, it can eventually settle in a state in which matter is so dense that electrons cannot remain bound to individual nuclei, since the distance between neighboring nuclei is smaller than the width of even the tightest bound electron wavefunction. This state can be described as an electron gas electrostatically bound to a compact fluid of nuclei kept together by gravitational attraction. There is no source of energy in such a star, and therefore its light is emitted at the expense of its internal kinetic energy. Stars in this state possess a high surface temperature, which can exceed 10^7 K (although the temperature decreases very slowly, since the emitted energy is not replenished), and a very small radius, of the order of a few thousand kilometers, comparable to earth's radius. Therefore their emitted light is bluish or white, but has a very small intensity.

Stars of this kind are called **white dwarfs** and, as we shall see, have masses comparable to that of the sun. Thus, since the star contains a mass of the order of a solar mass in a volume of the order of 1% of the solar radius, its density is of the order of a million times the density of the sun.

How can such a system avoid gravitational collapse? The force withstanding gravity is provided by the pression of the electron gas, due to the exclusion principle. In this section, following E. C. Stoner [Ston29], we treat a highly simplified model of a white dwarf, and show that there is an upper limit to its mass, which is of the order of (but larger than) the solar mass. The story of the discovery of this limit mass is nicely described in [Naue08].

We consider therefore a system made by N electrons, N protons, and N neutrons. Indeed, when the fission processes end, the star is essentially made of ^{4}He atoms, with an equal number of neutrons and protons. (In the subsequent processes leading to the white dwarf state, the nucleons rearrange into carbon and oxygen nuclei.) We assume that the Fermi energy of the electron is so large that we can treat it at zero temperature; this assumption has to be checked at the end of the calculation. Assuming a uniform density, the Fermi momentum p_{F} satisfies the relation

$$\frac{4}{3h^3}\pi p_{\mathrm{F}}^3 = \frac{N}{2V}, \tag{4.164}$$

where the factor 2 comes from the spin degeneracy. It is useful to express p_{F} as a function of the radius $R = (3V/4\pi)^{1/3}$ of the star:

$$p_{\mathrm{F}} = \left(\frac{9N}{32\pi^2}\right)^{1/3} \frac{h}{R} = \kappa N^{1/3} \frac{h}{R}, \tag{4.165}$$

where the constant κ is given by

$$\kappa = \left(\frac{9}{32\pi^2}\right)^{1/3} \simeq 3.054 \cdot 10^{-1}. \tag{4.166}$$

To evaluate the Fermi energy, we have to take into account the fact that the electron momentum can be so large that its speed is close to the speed of light. We need therefore to use the relativistic relation between energy and momentum:

$$\epsilon(p) = \sqrt{p^2c^2 + m^2c^4}, \tag{4.167}$$

where m is the electron mass. We have therefore the total kinetic energy of the electron gas, including its rest energy:

$$\begin{aligned} K &= \frac{2}{h^3} \int_0^{p_{\mathrm{F}}} \mathrm{d}p\, 4\pi p^2\, \epsilon(p) = \frac{32\pi^2\, m^4 c^5}{3h^3} R^3 \int_0^{x_{\mathrm{F}}} \mathrm{d}x\, x^2 \sqrt{1+x^2} \\ &= \frac{3m^4c^5}{\kappa^3 h^3} R^3 \mathcal{I}(x_{\mathrm{F}}), \end{aligned} \tag{4.168}$$

where we have defined

$$x_{\mathrm{F}} = \frac{p_{\mathrm{F}}}{mc}. \tag{4.169}$$

The integral $\mathcal{I}(x_{\mathrm{F}})$ appearing in eq. (4.168) can be calculated exactly, but it is more instructive to consider its limits:

- The classical limit, $x_F \ll 1$: it yields

$$\mathcal{I}(x_F) = \frac{x_F^3}{3}\left(1 + \frac{3x_F^2}{10} + \cdots\right). \tag{4.170}$$

- The ultrarelativistic limit, $x_F \gg 1$: it yields

$$\mathcal{I}(x_F) = \frac{x_F^4}{4}\left(1 + \frac{1}{x_F^2} + \cdots\right). \tag{4.171}$$

We now consider the gravitational energy of the system, which is essentially due to the nucleons, since the electron mass is so small. For simplicity, we assume that the total mass M is uniformly distributed in a sphere of radius R. Then it is known that the gravitational energy E_M is given by

$$E_M = -\frac{3}{5}G\frac{M^2}{R}. \tag{4.172}$$

We now need a relation between N and M. Since for each electron there are two nucleons, always neglecting the contribution of the electrons to the total mass, we obtain

$$N = \frac{M}{2m_n}, \tag{4.173}$$

where m_n is the nucleon mass. Summarizing, the total energy of the system can be expressed in terms of M and R by

$$E(M, R) = K + E_M = \frac{3m^4c^5}{\kappa^3 h^3}R^3\mathcal{I}\left(\frac{\kappa N^{1/3}h}{mcR}\right) - \frac{3}{5}G\frac{M^2}{R}, \tag{4.174}$$

where N is expressed in terms of M by eq. (4.173).

We now look for the minimum of the energy as a function of R, when M is given. In the classical limit $x_F \ll 1$ we obtain

$$E(M, R) = Nmc^2\left(1 + \frac{3}{10}\frac{\kappa^2 N^{2/3}h^2}{m^2c^2R^2}\right) - \frac{3}{5}G\frac{M^2}{R}. \tag{4.175}$$

This expression has a minimum for

$$R_{cl} = \frac{\kappa^2}{2^{5/3}}\frac{h^2}{mm_n^{5/3}}M^{-1/3}. \tag{4.176}$$

On the other hand, in the ultrarelativistic limit, $x_F \gg 1$, we have

$$E(M, R) = \frac{\alpha M^{4/3}}{R} + \beta M^{2/3}R - \frac{\gamma M^2}{R}, \tag{4.177}$$

where

$$\alpha = \frac{3\kappa hc}{8\cdot 2^{1/3}m_n^{4/3}}, \qquad \beta = \frac{3m^2c^3}{2\cdot 2^{2/3}\kappa h m_n^{2/3}}, \qquad \gamma = \frac{3}{5}G. \tag{4.178}$$

This expression has a minimum for $R = R_{rel}$, where

$$R_{rel} = \sqrt{\frac{\gamma M^2 - \alpha M^{4/3}}{\beta M^{2/3}}} = M^{1/3}\sqrt{\frac{\alpha}{\beta}}\sqrt{1 - \left(\frac{M}{M_0}\right)^{2/3}}, \tag{4.179}$$

in which we have defined

$$M_0 = \left(\frac{\alpha}{\gamma}\right)^{3/2} = \frac{\kappa^{3/2}}{4}\left(\frac{hc}{G}\right)^{3/2}\frac{1}{m_n^2}. \tag{4.180}$$

Clearly, there is no solution for $M > M_0$. This implies that a star whose mass exceeds M_0 cannot resist collapsing gravitationally either into an even denser state, or into a black hole. Our result, assuming a uniform density inside the star, yields $M_0 \simeq 2.448 \cdot 10^{30}$ kg, which corresponds to roughly 1.232 solar masses. A more careful calculation, allowing for a nonuniform mass distribution, yields the **Chandrasekhar limit** [Chan31] $M_C \simeq 2.765 \cdot 10^{30}$ kg, corresponding to about 1.391 solar masses. In terms of M_0 the relation between R and M reads

$$R = \sqrt{\frac{\gamma}{\beta}} M^{2/3}\left(\frac{M_0}{M}\right)^{1/3}\sqrt{1-\left(\frac{M}{M_0}\right)^{2/3}}. \tag{4.181}$$

The factor $\sqrt{\gamma/\beta}$ is approximately equal to $3.376 \cdot 10^{-14}$ in SI units. The mass ratio M/M_0 is of order 1. Thus the radius R is of the order of 10^3 km, but decreases to 0 as $M \to M_0$.

Exercise 4.10 We consider a typical white dwarf of mass $M = 2.0 \cdot 10^{30}$ kg.

1. Evaluate the Fermi momentum p_F of the electrons in the star. Compare its kinetic energy $p_F c$ with its rest energy mc^2. Is the ultrarelativistic approximation warranted?
2. Evaluate the Fermi temperature of the electrons in the white dwarf and compare it with the typical value of 10^7 K.
3. Evaluate the pressure due to the nuclei, considered as a perfect non-relativistic gas, and compare it to the pressure of the electron gas. (One can take equal ratios of oxygen and carbon.)

Exercise 4.11 (Neutron stars) Massive stars (whose mass exceeds about ten solar masses) avoid the white dwarf fate, because their gravitational energy is large enough to allow the continuation of nuclear synthesis even if all hydrogen is transformed into helium. Then, after a number of processes that we do not attempt to summarize here, the star relinquishes a ball entirely made of neutrons, with mass about 1.4 times the solar mass, kept together by their gravitational attraction. This state is known as a **neutron star**. Here we assume that the neutron gas is a nonrelativistic Fermi gas at zero temperature.

1. Given the mass M of the star and its radius R, give the expression of the internal kinetic energy of the star.
2. Give the expression of the total energy of the star, including its gravitational contribution, and assuming that its density is uniform. Evaluate its equilibrium radius as a function of M.
3. Evaluate the radius R for a star with mass M of the order of 1.4 times the solar mass. Evaluate the Fermi temperature of the neutrons and compare it with the initial temperature of the star. Is the zero-temperature approximation warranted?

4. Evaluate the escape speed of an object of mass m placed at the surface of the neutron star. Evaluate the value M_{BH} of a star for which this speed equals the speed of light.

4.8 Variational Derivation of Fermi and Bose Statistics

In the present section, we show that it is possible to obtain the Fermi and Bose distributions (4.92) and (4.94) from a variational principle on the quantity $\mathcal{S} = \ln \mathcal{N}$, where $\mathcal{N}$ is the number of allowed microscopic configurations, that is, the number of different ways in which we can assign occupation numbers to single-particle states, keeping the total number of particles and the total energy constant. In other words, we evaluate $\mathcal{N}$ for each distribution of the occupation numbers that satisfies the constraints on the total energy and the total number of particles. We then identify the equilibrium distribution as the one that corresponds to the maximal value of $\mathcal{N}$.

Fermi statistics

Let us suppose that we have a system of N fermions, each of which can be in a single-particle state i of energy ϵ_i. We want to evaluate the average value n_i of the occupation number of state i. Although this problem can be resolved more easily in the grand canonical ensemble, it is instructive to obtain it by using the variational principle for the generalized ensembles. In this case, the microstate of the system is univocally defined by the collection of the occupied states $\{i_1, \ldots, i_N\}$. Since the probability of occupation of a state ultimately depends only on its energy ϵ_i, we can collect the single-particle states in groups such that the number of states G_j that belong to group j is large, the number of particles N_j that are found in group j is also large, but the energies of all the states that belong to the same group are approximately equal. The ensemble is therefore well defined when, for each group j, its size G_j, the corresponding energy ϵ_j, and the average occupation number $n_j = N_j/G_j$ are all given. Once N_j is fixed, the number of ways in which we can distribute the particles within group j is clearly given by $G_j!/\left(N_j!(G_j - N_j)!\right)$, which corresponds to the number of ways in which it is possible to choose the N_j states that are occupied from among all the G_j possible ones. The logarithm of this quantity is given by

$$s_j = -G_j \left[n_j \ln n_j + (1 - n_j) \ln(1 - n_j) \right], \tag{4.182}$$

where we have taken advantage of Stirling's formula.

We thus obtain

$$\mathcal{S} = \sum_j s_j = -\sum_j G_j \left[n_j \ln n_j + (1 - n_j) \ln(1 - n_j) \right]. \tag{4.183}$$

We look for the maximum of this quantity, subject to the constraints

$$\sum_j N_j = \sum_j G_j n_j = N; \qquad \sum_j G_j \epsilon_j n_j = E. \tag{4.184}$$

By introducing the Lagrange multipliers α and β, we look for the extremum of

$$\mathfrak{S} = \mathcal{S} + \alpha \sum_j G_j n_j + \beta \sum_j G_j \epsilon_j n_j. \tag{4.185}$$

We thus obtain the equation

$$-\ln \frac{n_j}{1-n_j} = \alpha + \beta \epsilon_j, \tag{4.186}$$

which corresponds to

$$n_j = \frac{1}{\mathrm{e}^{\alpha+\beta\epsilon_j}+1}. \tag{4.187}$$

This is the Fermi distribution, expressed as a function of the parameters $\alpha = -\mu/k_\mathrm{B}T$ and $\beta = 1/k_\mathrm{B}T$.

Bose statistics

Let us now consider a system of N bosons, and let us proceed once again to grouping the single-particle states together. In this case, the number of ways in which it is possible to arrange the N_j particles in the G_j states is given by the number of ways in which the N_j states can be chosen *with repetition*. This number can be evaluated by considering that it is equal to the number of ways one can arrange N_j indistinguishable marbles in G_j distinguishable boxes. We represent an arrangement of this sort in the following manner:

$$|\bullet\bullet|\bullet||\bullet\bullet\bullet| \qquad \text{etc.},$$

where the vertical bars sequentially mark the borders of the first, the second, …, the nth box, and the dots represent the marbles. We have N_j marbles and $G_j + 1$ bars in total, but the positions of the first and last bars are fixed, at the beginning and end of the sequence, respectively. The number of arrangements is therefore equal to the number of ways in which we can place the N_j marbles in $N_j + G_j - 1$ possible positions. This number is equal to $(N_j + G_j - 1)!/\left(N_j!(G_j-1)!\right)$, and its logarithm is given by

$$s_j = G_j\left[(1+n_j)\ln(1+n_j) - n_j \ln n_j\right], \tag{4.188}$$

where $n_j = N_j/G_j$ is the average occupation number, and we have assumed $G_j \gg 1$. We then look for the extremum of

$$\mathcal{S} = \sum_j s_j = k_\mathrm{B} \sum_j G_j\left[(1+n_j)\ln(1+n_j) - n_j \ln n_j\right], \tag{4.189}$$

subject to the same constraints as above. We obtain

$$n_j = \frac{1}{\mathrm{e}^{\alpha+\beta\epsilon_j}-1}. \tag{4.190}$$

By interpreting the parameters α and β as we did earlier, we obtain the Bose distribution

$$n_j = \frac{1}{e^{(\epsilon_j - \mu)/k_B T} - 1}, \tag{4.191}$$

where we assumed

$$\mu \leq \min_j \epsilon_j, \tag{4.192}$$

because the n_js must not be negative.

4.9 Einstein Condensation

Let us now consider the behavior of a strongly degenerate boson gas. We analyze the expression that we derived for the average number of particles as a function of z, and we recall that since $\mu < 0$, one has $z < 1$. For generality's sake, let us assume that we have a system in a d-dimensional space. We thus have, by generalizing eq. (4.90),

$$\begin{aligned} N &= \frac{V}{\lambda_B^d}(2s+1)\frac{S_d}{\pi^{d/2}}\int_0^\infty dx\, x^{d-1}\frac{1}{e^{x^2}/z - 1} \\ &= \frac{V}{\lambda_B^3}(2s+1)\phi_{d/2}(z), \end{aligned} \tag{4.193}$$

where $\phi_\alpha(z) = \sum_n z^n/n^\alpha$ is the polylog function. The integral can be evaluated by the method shown in eq. (4.87). Since $\phi_\alpha(z)$ is an increasing function of z, it assumes its maximum value at $z = 1$. For $d \leq 2$, this maximum value is infinite but for $d = 3$, for example, one has

$$\phi_{3/2}(z) \leq \phi_{3/2}(1) = \zeta_R\left(\frac{3}{2}\right) = 2.612\ldots. \tag{4.194}$$

It therefore seems that the quantity $N\lambda_B^3/V$ cannot be larger than a certain value. On the other hand, when T decreases, the thermal de Broglie wavelength increases, and therefore, for a system with N particles contained in a fixed volume, this quantity becomes arbitrarily large at low temperatures. Einstein recognized this problem while reading an article sent to him by an obscure Bengali physicist, Satyendranath Bose, asking for his approval for publication, and he then suggested the solution (see [Bose24] and [Eins25]).

Let us recall that approximating the sum with the integral when expressing N is allowed only if each term in the sum gives a negligible contribution to the total. So if $\mu \to 0$ (in other words, if $z \to 1$), the Bose factor for the lowest single-particle energy values—in particular, for $\epsilon(\boldsymbol{p}=0)$—becomes very large (as can be seen from the plot of the Bose factor that diverges for $x \to 0$). It is therefore useful to separate the contribution of the term $\boldsymbol{p} = 0$ and calculate the corresponding value of z:

$$N = (2s+1)\left[\frac{z}{1-z} + \frac{V}{\lambda_B^d}\phi_{d/2}(z)\right]. \tag{4.195}$$

The first term is usually negligible when compared to the rest, which is proportional to V, but since the second term cannot be larger than $(V/\lambda_B^d)\phi_{d/2}(1)$, one needs to account for it if N becomes too large (or T too small). In this situation, we see that one has to choose a value of z that is very close to 1, and it is therefore possible to make the approximation

$$N \simeq (2s+1)\left[\frac{1}{1-z} + \frac{V}{\lambda_B^d}\phi_{d/2}(1)\right] = N_0 + (N - N_0). \tag{4.196}$$

We thus obtain the occupation number N_0 of the single-particle state at $\boldsymbol{p} = 0$:

$$N_0 = N - \frac{V(2s+1)}{\lambda_B^d}\phi_{d/2}(1) = N\left[1 - \frac{2s+1}{\rho}\left(\frac{2\pi m k_B T}{h^2}\right)^{d/2}\phi_{d/2}(1)\right], \tag{4.197}$$

where $\rho = N/V$ is the particle density. We also have

$$1 - z = \frac{2s+1}{N_0}. \tag{4.198}$$

Clearly, these expressions make sense only for $N_0 > 0$—in other words, if

$$T < T_c = \frac{h^2}{2\pi m k_B}\left(\frac{\rho}{\phi_{d/2}(1)(2s+1)}\right)^{2/d}. \tag{4.199}$$

This defines the **condensation temperature** T_c. For $T < T_c$, the number of particles that occupy the minimum energy state is macroscopic: one says that the particles have **condensed**. This phenomenon is called the **Einstein condensation** (sometimes, illegitimately, Bose-Einstein condensation).

Below the transition temperature, the number of particles in a state of minimum energy ("in the condensate") is given by

$$N_0 = N\left[1 - \left(\frac{T}{T_c}\right)^{2/d}\right], \tag{4.200}$$

and the pressure depends only on the temperature:

$$p = \frac{k_B T}{\lambda_B^d}\phi_{(d+2)/2}(1) = \frac{k_B T}{\lambda_B^d}\zeta_R\left(\frac{d+2}{2}\right) \propto T^{(d+2)/2}. \tag{4.201}$$

In particular, in three dimensions one has $p = k_B T \zeta_R(5/2)/\lambda_B^3 \simeq 1.342\, k_B T/\lambda_B^3$.

The equation of state of an ideal Bose gas in three dimension is shown in fig. 4.7. When the parameter $\rho\lambda_B^3/(2s+1)$ exceeds $\zeta_R(3/2)$, the quantity $p\lambda_B^3/((2s+1)k_B T)$ remains constant at the value $\zeta_R(5/2)$. The condensation point is marked by an arrow in the plot.

The phenomenon of Einstein condensation is the basis for the macroscopic quantum properties of helium (superfluidity) and of superconductivity. In these cases, however, the interactions among particles are very important, to the extent that they qualitatively change the phenomenon. We cannot therefore discuss them here.

We shall conclude by observing that if $d \leq 2$, the Einstein condensation does not take place—in this case, in fact, $\lim_{z\to 1}\phi_{d/2}(x) = \infty$. Therefore, for any value of N, it is possible to find a value of $z < 1$, and the contribution of the minimum energy state is negligible.

Exercise 4.12 Evaluate the behavior of the internal energy and specific heat of a boson gas in the vicinity of the Einstein condensation.

Exercise 4.13 (Einstein condensation of zero-mass particles) Consider a system of quantum particles with spin zero, whose energy, as a function of the wave vector $\boldsymbol{k}$, is given by

$$\epsilon(\boldsymbol{k}) = c|\boldsymbol{k}|.$$

The system is contained in a cubic d-dimensional box of volume $V = L^d$.

1. Give the expression of the system's grand canonical partition function.
2. Deduce the expression for the average number N of particles.
3. Obtain, as a function of V, c, and T, the value N_c of the particle number in which Einsten's condensation takes place. Does it take place for $d = 2$? And for $d = 1$?

(It is not necessary to evaluate explicitly *convergent* integrals.)

Exercise 4.14 (Einstein condensation in a harmonic potential) Consider a system of N quantum particles of spin zero and mass m in d dimensions, subject to a harmonic potential of the form

$$u(\boldsymbol{r}) = \frac{1}{2} m\omega_0^2 r^2, \tag{4.202}$$

where $\boldsymbol{r} = (r_1, \ldots, r_d)$ is the position vector, and ω_0 is a parameter with the dimensions of a frequency.

1. Give the expression of the grand canonical partition function and show that the chemical potential cannot exceed a special value μ_0. Give the expression of μ_0.
2. Give the expression of the number $N - N_0$ of particles in the excited states. Approximate the resulting expression by an integral for $d = 2$, $d = 3$.
3. Define $R_0 = \sqrt{u_0/m}/\omega_0$, where u_0 is a parameter with the dimensions of energy. The thermodynamic limit corresponds to $N \to \infty$, $N/R_0^d = \rho = \text{const.}$, which implies $\omega_0 \to 0$. Show that in this limit the system exhibits an Einstein condensation for $d = 2$ or $d = 3$ and give the expression of the corresponding transition temperature.

4.10 Adsorption

Let us consider a wall that can adsorb a species of molecules placed in contact with a container that holds those same molecules in a gaseous state. To say that the wall can adsorb molecules is the equivalent of saying that their energy when they are on the wall is lower than when they are in the gas. Let us denote by $-\epsilon_0$ ($\epsilon_0 > 0$) the energy of an absorbed molecule. On the other hand, the wall's capacity is finite. Let us assume that there can, at

most, be only N particles deposited on it. When a molecule has settled on the wall, another molecule cannot adhere to it at the same point.

These observations lead to the introduction of the following Hamiltonian to describe the state of the wall:

- One introduces a variable $\tau_i \in \{0, 1\}$ for each adsorption site: 0 if the site is empty, 1 if the site is full.
- The energy of the adsorbed molecules is given by $\mathcal{H}(\tau) = -\epsilon_0 \sum_{i=1}^{N} \tau_i$, where the microstate τ is identified by the collection of the τ_i: $\tau = (\tau_i)$.
- Since the gas can freely exchange particles with the surrounding gas, the number of adsorbed particles is not fixed—we must, in other words, use the grand canonical ensemble.

The partition function is therefore given by

$$Z = \sum_{\tau} \exp\left(\frac{1}{k_B T} \sum_{i=1}^{N} \tau_i(\epsilon_0 + \mu)\right), \tag{4.203}$$

and recalls that of a fermion system. By evaluating the sum, we obtain

$$\ln Z = N \ln\left[1 + \exp\left(\frac{\epsilon_0 + \mu}{k_B T}\right)\right]. \tag{4.204}$$

The average number of adsorbed particles is given by

$$\langle M \rangle = \frac{\partial \ln Z}{\partial(\mu / k_B T)} = N\left[\exp\left(-\frac{\epsilon_0 + \mu}{k_B T}\right) + 1\right]^{-1}. \tag{4.205}$$

This expression can be related to the pressure of the gas surrounding the wall by exploiting the fact that, for a classical ideal gas, one has

$$\exp\left(\frac{\mu}{k_B T}\right) = z = \frac{p \lambda_B^3}{k_B T}. \tag{4.206}$$

We thus obtain

$$\langle M \rangle = N \frac{p}{p + p_0}, \tag{4.207}$$

where we have introduced the characteristic pressure p_0 by means of the expression

$$\exp\left(-\frac{\epsilon_0}{k_B T}\right) = \frac{p_0 \lambda_B^3}{k_B T}. \tag{4.208}$$

We observe that the characteristic pressure p_0 corresponds to 50% of occupied sites. This law is known as the **Langmuir isotherm** and its behavior is shown in figure 4.10, where we plot $\langle M \rangle / N$ vs. p/p_0. We have $M/N \simeq p/p_0$ for $p \ll p_0$ and $\lim_{p \to \infty} M/N = 1$.

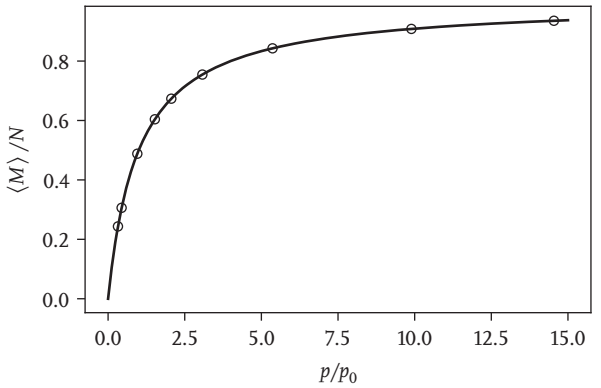


FIGURE 4.10. The Langmuir isotherm. The average occupation number per site, $\langle M \rangle / N$, is plotted vs. the rescaled pressure, p/p_0. Circles denote Langmuir's data on adsorption of nitrogen on mica [Lang18].

4.10.1 Myoglobin and Hemoglobin

The Langmuir isotherm has important applications in biophysics. Some proteins specialize in transporting oxygen—the simplest is myoglobin, which is also the first protein whose three-dimensional structure was solved. Myoglobin can be found in muscles, and it is responsible for their characteristic red color. In one of the protein's "pockets" there is a complex organic group, called **heme**, which also includes an iron ion, which is responsible for the protein's activity. A set of N proteins behaves like the adsorbing wall that we just discussed. The plot of the average number of oxygen molecules adsorbed by myoglobin as a function of the partial pressure of oxygen has a typical Langmuir trend like that in figure 4.10.

Hemoglobin is another protein, present in the blood, whose task is to carry oxygen from the lungs to the muscles. It is composed of four units, each with a structure similar to myoglobin's. It is important that hemoglobin adsorbs oxygen where the partial pressure of oxygen is highest and release it where it is lowest. One can observe that the corresponding isotherm has a fairly different behavior from that of myoglobin at low oxygen pressures—this feature, due to the cooperative effort of the four units, increases the protein's efficiency. The characteristic pressure of hemoglobin, moreover, decreases as pH levels become lower, especially when muscles are tired—this is the Bohr effect, named for Niels Bohr's father, who was a famous physiologist.

The fetus hemoglobin has a smaller characteristic pressure than that of an adult, so it can take oxygen from the mother. In certain wild geese who migrate above the Himalayas, a single mutation in an amino acid significantly decreases p_0 so as to increase the oxygen adsorbed at great altitudes.

Exercise 4.15 (A simplified model of hemoglobin) Let us associate each of the four units i ($i = 1, \ldots, 4$) that make up a molecule α of hemoglobin with a variable $\tau_{\alpha i}$, which is equal to 1 if an O_2 molecule is adsorbed, and vanishes otherwise. The energy of the hemoglobin molecule is given by

- $-\epsilon_0$ times the number of O_2 molecules adsorbed and
- $-J$ times the number of pairs of nearest-neighbor units, each containing a molecule of adsorbed O_2. (The i and $i+1$ sites are nearest neighbors, and so are sites 4 and 1).

Consider a system of $N/4$ molecules of hemoglobin. Evaluate the average number $M = \sum_{\alpha=1}^{N/4} \sum_{i=1}^{4} \langle \tau_{\alpha i} \rangle$ of adsorbed molecules as a function of $x = \exp[(\epsilon_0 + \mu)/k_B T]$ and of $y = \exp(J/k_B T)$.

4.11 Internal Degrees of Freedom

Up to now, we have assumed that the particles being considered have no internal structure. The case in which each particle has a certain number of internal degrees of freedom can be treated in a simple fashion, as long as we can disregard the interactions between the different particles.

4.11.1 Atomic Gases

The first case of interest is that of a diluted atomic gas. In this case, in addition to the translational degrees of freedom, one also needs to consider the effects of the electronic and nuclear degrees of freedom. These degrees of freedom are independent of the translational ones (which are those we have treated so far), with excellent approximation. We can therefore identify the state of an atom by means of the quantity

$$\nu = (\boldsymbol{p}, n, \sigma), \tag{4.209}$$

where $\boldsymbol{p}$ is momentum and defines the translational state of the center of mass, n denotes the electronic state of the atom, and σ denotes the state of the nucleus.

The energy of the atom in the state $\nu = (\boldsymbol{p}, n, \sigma)$ is given by

$$\epsilon(\nu) = \frac{p^2}{2m} + \epsilon_{n,\sigma}, \tag{4.210}$$

which expresses the independence of the internal degrees of freedom.

Assuming that there are no interactions among different particles, we can evaluate the canonical partition function by summing over the individual degrees of freedom of each particle independently:

$$\begin{aligned} Z &= \frac{1}{N!}\left[\int \frac{\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{r}}{h^3} \sum_{n,\sigma} \mathrm{e}^{-\epsilon(\boldsymbol{p},n,\sigma)/k_B T}\right]^N \\ &= \frac{1}{N!}\left[\int \frac{\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{r}}{h^3} \mathrm{e}^{-p^2/2mk_B T}\right]^N \left[\sum_{n,\sigma} \mathrm{e}^{-\epsilon_{n,\sigma}/k_B T}\right]^N \\ &= \frac{1}{N!} \zeta_{\mathrm{transl}}^N(T, V) \zeta_{\mathrm{int}}^N(T). \end{aligned} \tag{4.211}$$

Here we introduced the single-particle **translational** and **internal** partition functions, defined by

$$\zeta_{\text{transl}}(T, V) = \int \frac{\mathrm{d}\boldsymbol{p}\,\mathrm{d}\boldsymbol{r}}{h^3} \exp\left(-\frac{p^2}{2mk_{\text{B}}T}\right) = \frac{V}{\lambda_{\text{B}}^3}; \tag{4.212}$$

$$\zeta_{\text{int}}(T) = \sum_{n,\sigma} \exp\left(-\frac{\epsilon_{n,\sigma}}{k_{\text{B}}T}\right) = \exp\left(-\frac{\epsilon_0}{k_{\text{B}}T}\right) \sum_{n,\sigma} \exp\left(-\frac{\epsilon_{n,\sigma} - \epsilon_0}{k_{\text{B}}T}\right). \tag{4.213}$$

In the second equation, ϵ_0 is the ground-state energy of the atom.

The translational partition function is the one we examined in the preceding sections. There we evaluated the corresponding partition function classically, which is usually acceptable at a high enough temperature. In calculating the internal partition function, one has to keep in mind that the differences in energy that appear in its expression are usually much larger than $k_{\text{B}}T$ (in order of magnitude, 1 eV corresponds to about 10^4 K). It is therefore necessary to consider only those states that have the same energy as the ground state, and we then obtain

$$\begin{aligned}\zeta_{\text{int}}(T) &\simeq \exp\left(-\frac{\epsilon_0}{k_{\text{B}}T}\right) \times (\text{ground-state multiplicity}) \\ &= \exp\left(-\frac{\epsilon_0}{k_{\text{B}}T}\right) g_{\text{el}}^0 g_{\text{nucl}}^0,\end{aligned} \tag{4.214}$$

where g_{el}^0 and g_{nucl}^0 are the multiplicities of the electronic and nuclear ground states, respectively. For example, if the only important nuclear degree of freedom of the nucleus is its spin, we obtain

$$g_{\text{nucl}}^0 = 2I + 1, \tag{4.215}$$

where I is the magnitude of the nuclear spin. We thus obtain the expression for the canonical partition function and, consequently, that for the Helmholtz free energy:

$$\ln Z = N \ln\left[\frac{\mathrm{e}V}{\lambda_{\text{B}}^3}\right] + N \ln\left[(2I+1)g_{\text{el}}^0\right] - \frac{N\epsilon_0}{k_{\text{B}}T}. \tag{4.216}$$

Since the dependence on volume is entirely contained in the translational partition function, we obtain that the pressure (and therefore the equation of the isotherms) is independent of the internal degrees of freedom, which do, however, influence the value of the entropy and the internal energy.

Exercise 4.16 (Particles with variable spin) Let us consider a system made up of N particles, each of which can be in a spin $s = 0$ state (with vanishing energy) or in one of three spin $s = 1$ states, where the z component of the spin is equal to $\sigma \in \{-1, 0, +1\}$, with energy equal to $\epsilon_0 - h\sigma$, where h is the applied magnetic field, measured in suitable units.

1. Evaluate the canonical partition function of the system at the absolute temperature T. (Hint: Introduce the variable N_1, equal to the number of particles that are in the $s = 1$ state).

2. Provide the expressions of the internal energy E and of the magnetization $M = \left\langle \sum_i \sigma_i \right\rangle$ as functions of T and h.

3. Discuss the behavior of the susceptibility per particle $\chi = \partial M / \partial h)_{h=0} / N$ as a function of the temperature.

Exercise 4.17 (Frustrated spins) A system is made up of N Ising spins, $\sigma_i \in \{-1, +1\}$, $i = 1, \ldots, N$, that interact in groups of three, with Hamitonian

$$\mathcal{H}_\ell^{(3)} = J\left(\sigma_{3\ell-1}\sigma_{3\ell-2} + \sigma_{3\ell-2}\sigma_{3\ell} + \sigma_{3\ell}\sigma_{3\ell-1}\right) - h\sum_{i=0}^{2}\sigma_{3\ell-i},$$

$$\ell = 1, \ldots, N/3,$$

with a *positive* interaction coefficient J. Thus the total Hamiltonian of the system is given by

$$\mathcal{H} = \sum_{\ell=1}^{N/3} \mathcal{H}_\ell^{(3)}.$$

1. Evaluate the possible energy values of each "elementary triangle" and its corresponding degeneration.

2. Evaluate the equation of state for the magnetization per spin $m = \langle \sigma_i \rangle$,

   $$m = m(h, T),$$

   and discuss its behavior.

4.11.2 Molecular Gases

Let us now consider a gas composed of molecules of a specific chemical species. The situation becomes more complicated for two reasons:

- One has to take into account the rotational degrees of freedom of the molecules, which have characteristic energies that are comparable to $k_B T$.
- The degrees of freedom for both the electronic and the nuclear spins are not independent of the rotational ones in the case of molecules composed of identical atoms, because of the Pauli principle.

First, let us attempt an accurate description of the quantum behavior of a molecule. Strictly speaking, we ought to consider the whole problem of the interaction between nuclei and electrons. However, since the ratios of the masses of the nuclei to those of the electrons are large, the characteristic speeds of the nuclei are much smaller than those of the electrons. We can the introduce the **Born-Oppenheimer approximation**, which amounts to supposing that the positions of the nuclei are fixed, which allows us to solve the Schrödinger equation for the electron degrees of freedom in a fixed external potential. We thus obtain a fixed value for energy for the given fixed positions of the nuclei. This

energy is then considered as an effective potential in the Schrödinger equation for the nuclei.

More explicitly, let us express the total wave function of the system (nuclei + electrons) in the following way:

$$\Psi(r, R) = \Psi_n^{(\mathrm{el})}(r, R)\Psi_{n,\nu}^{(\mathrm{nucl})}(R), \tag{4.217}$$

where $r = (\boldsymbol{r}_1, \ldots \boldsymbol{r}_{N\mathrm{el}})$ denotes the electronic coordinates and $R = (\boldsymbol{R}_1, \ldots, \boldsymbol{R}_{N_{\mathrm{nucl}}})$ the coordinates of the nuclei, and we denote respectively by n and ν the quantum numbers of the electrons and the nuclei. The electronic wave function $\Psi_n^{(\mathrm{el})}(r, R)$ satisfies the equation

$$\left[\widehat{K}(r) + V(r) + V(R)\right] \Psi_n^{(\mathrm{el})}(r, R) = \epsilon_n(R)\Psi_n^{(\mathrm{el})}(r, R), \tag{4.218}$$

where $\widehat{K}(r)$ is the kinetic energy operator for the electrons only. The wave function of the nuclei satisfies the equation

$$\left[\widehat{K}(R) + \epsilon_n(R)\right] \Psi_{n,\nu}^{(\mathrm{nucl})}(R) = \epsilon_{n,\nu}\Psi_{n,\nu}^{(\mathrm{nucl})}(R), \tag{4.219}$$

where $\widehat{K}(R)$ is the kinetic energy operator for the nuclei. In the context of this approximation, we have in general

$$\zeta_{\mathrm{int}}(T) = \sum_{n,\nu} \exp\left(-\frac{\epsilon_{n,\nu}}{k_{\mathrm{B}}T}\right). \tag{4.220}$$

As we previously saw, it is usually possible to disregard all the excited electronic states, since the energies that come into play are much larger than $k_{\mathrm{B}}T$. We therefore obtain

$$\begin{aligned}\zeta_{\mathrm{int}}(T) &\simeq \sum_{\nu} \exp\left(-\frac{\epsilon_{0,\nu}}{k_{\mathrm{B}}T}\right) \\ &= g^0 \exp\left(\frac{\epsilon_{0,0}}{k_{\mathrm{B}}T}\right) \sum_{\nu} \exp\left(-\frac{\epsilon_{0,\nu} - \epsilon_{0,0}}{k_{\mathrm{B}}T}\right),\end{aligned} \tag{4.221}$$

where g^0 is the total multiplicity of the ground state of the molecule.

It is usually possible to consider the vibrational and rotational degrees of freedom as independent, because the vibrations take place much more rapidly than the rotations. We thus obtain

$$\Delta\epsilon_\nu = \epsilon_{0,\nu} - \epsilon_{0,0} = \hbar\omega_0\left(n + \frac{1}{2}\right) + \frac{\hbar^2}{2\mathcal{I}}\ell(\ell+1). \tag{4.222}$$

Here, $n = 0, 1, 2, \ldots$ is the quantum vibration number, and ω_0 is the vibration frequency around the minimum of the Born-Oppenheimer potential, which corresponds to a distance R_0 between nuclei:

$$\omega_0^2 = \left(\frac{1}{m_{\mathrm{A}}} + \frac{1}{m_{\mathrm{B}}}\right) \left.\frac{\partial^2\epsilon_0(R)}{\partial R^2}\right|_{R=R_0}. \tag{4.223}$$

The first factor is the inverse reduced mass of the two nuclei. The second term in eq. (4.222) is the rotational energy, where $\ell = 0, 1, 2, \ldots$ is the angular quantum number, and $\mathcal{I}$ is the moment of inertia. The level identified by ℓ is $(2\ell + 1)$ times degenerate.

If the molecule is composed of different atoms, we can sum independently over the vibrational and rotational degrees of freedom. If we are dealing with identical atoms, we must be careful to sum only over those states that are either even or odd with respect to nuclear exchange, depending on whether the nuclei are bosons or fermions, respectively. In the classical limit, it is sufficient to divide by a symmetry factor that takes into account the double counting. If we assume that we can perform this approximation, we obtain

$$\zeta_{\text{int}}(T) = \frac{g^0}{\sigma_{\text{AB}}} \exp\left(-\frac{\epsilon_{0,0}}{k_{\text{B}}T}\right) (2I_{\text{A}} + 1)\,(2I_{\text{B}} + 1)\zeta_{\text{vib}}(T)\zeta_{\text{rot}}(T), \tag{4.224}$$

where I_{A} and I_{B} are the respective magnitudes of the nuclear spins for nuclei A and B, and the symmetry factor σ_{AB} is given by

$$\sigma_{\text{AB}} = \begin{cases} 1, & \text{if A} \neq \text{B}; \\ 2, & \text{if A} = \text{B}. \end{cases} \tag{4.225}$$

Calculating the vibrational partition function is simple:

$$\begin{aligned} \zeta_{\text{vib}}(T) &= \sum_{n=0}^{\infty} \exp\left[-\frac{\hbar\omega_0}{k_{\text{B}}T}\left(n + \frac{1}{2}\right)\right] \\ &= \exp\left(-\frac{\hbar\omega_0}{2k_{\text{B}}T}\right)\left[1 - \exp\left(-\frac{\hbar\omega_0}{k_{\text{B}}T}\right)\right]^{-1}. \end{aligned} \tag{4.226}$$

We can now calculate the contribution of the rotational degrees of freedom. As we saw, the quantum states of a rotator are specified by two integers: the **angular** quantum number ℓ, which can be equal to $0, 1, 2, \ldots$, and the **magnetic** quantum number m, which can be equal to $-\ell, -\ell + 1, \ldots, \ell - 1, \ell$. The kinetic energy of a quantum rotator in the (ℓ, m) state is given by $\epsilon(\ell, m) = \hbar^2\ell(\ell + 1)/2\mathcal{I}$, where $\mathcal{I}$ is the moment of inertia of the rotator. This energy does not depend on the magnetic quantum number, and for this reason, the energy levels are $(2\ell + 1)$ times degenerate.

In this fashion, we can evaluate the rotational partition function $\zeta_{\text{rot}}(T)$:

$$\zeta_{\text{rot}}(T) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{+\ell} \exp\left[-\frac{\hbar^2}{2\mathcal{I}}\ell(\ell + 1)\right] = \sum_{\ell=0}^{\infty} (2\ell + 1) \exp\left[-\frac{\hbar^2}{2\mathcal{I}}\ell(\ell + 1)\right]. \tag{4.227}$$

This expression cannot be represented in terms of elementary functions, but it is easy to calculate it numerically and also to obtain its behavior for large and small temperature values. We see that the characteristic temperature T_{rot} for rotational degrees of freedom is given by

$$T_{\text{rot}} = \frac{\hbar^2}{2\mathcal{I}k_{\text{B}}}. \tag{4.228}$$

In figure 4.11, we show the behavior of the contribution of rotational degrees of freedom to the specific heat—in other words, C_{rot}/N (in units of k_{B}) as a function of T/T_{rot}. We see

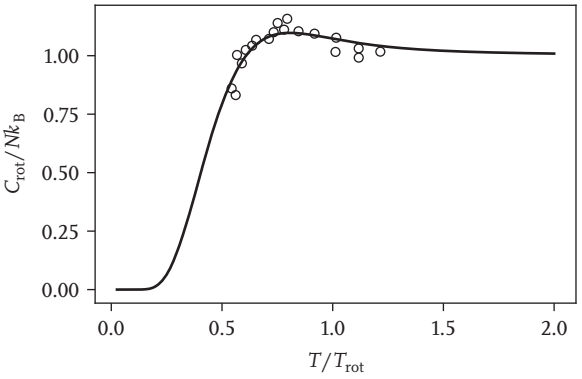


FIGURE 4.11. Rotational specific heat per molecule, as a function of T/T_{rot}, where $T_{\mathrm{rot}} = \hbar^2/2\mathcal{I}k_{\mathrm{B}}$. Data for the HD molecule from [Clus34].

that asymptotically $C_{\mathrm{rot}} \to 0$ for $T \ll T_{\mathrm{rot}}$. The curve exhibits a characteristic bump around $T_{\mathrm{rot}}/2$. The limit value at high temperatures is k_{B}, yielding a specific heat per molecule of diatomic gases equal to $5k_{\mathrm{B}}/2$. The fact that the value was different from $3k_{\mathrm{B}}$, as one would have taking into account two three-dimensional degrees of freedom per molecule, was of grave concern at the turn of the twentieth century. In fact, this is one of the puzzles that led Gibbs to state that [Gibb02, p. viii]

> we do not escape difficulties in as simple a matter as the number of degrees of freedom of a diatomic gas. It is well known that while theory would assign to the gas six degrees of freedom per molecule, in our experiments on specific heat we cannot account for more than five. Certainly, one is building on an insecure foundation, who rests his work on hypotheses concerning the constitution of matter.

Since T_{rot} is usually quite a bit smaller than the ambient temperature, it is worthwhile to expand $\zeta_{\mathrm{rot}}(T)$ in a power series of T_{rot}/T. This can be obtained by exploiting the Euler-MacLaurin series to evaluate the expression of ζ_{rot}:

$$\sum_{\ell=0}^{\infty} f(\ell) = \int_0^{\infty} \mathrm{d}\ell\, f(\ell) + \frac{1}{2} f(0) - \frac{1}{12} f'(0) + \frac{1}{720} f'''(0) + \cdots . \tag{4.229}$$

The result is

$$\zeta_{\mathrm{rot}}(T) = \frac{T}{T_{\mathrm{rot}}} \left[1 + \frac{1}{3} \left(\frac{T_{\mathrm{rot}}}{T} \right) + \frac{1}{15} \left(\frac{T_{\mathrm{rot}}}{T} \right)^2 + \cdots \right]. \tag{4.230}$$

Notice that the data shown in fig. 4.11 correspond to the hydrogen-deuterium (HD) molecule, rather than to the more common H_2 molecule. Indeed, the H_2 molecule has a peculiar behavior, which we discuss in the next subsection.

4.11.3 Ortho-Hydrogen and Para-Hydrogen

The effects of the quantum exclusion principle that we alluded to are particularly obvious in the case of hydrogen. In this case, moreover, it is necessary to discuss some of the subtleties of thermodynamic equilibrium fairly carefully.

Let us consider the H_2 molecule. The two nuclei are protons (with a spin $1/2$, and therefore fermions). In the Born-Oppenheimer approximation, we can consider only the wave function of the two nuclei (in an appropriate electronic state) and factorize the dependence on the motion of the center of mass. The state of the two nuclei is therefore described by a wave function that is a function of the two spins and of the relative position vector $\boldsymbol{R}$ of the two nuclei:

$$\Psi_{\sigma_1\sigma_2}(\boldsymbol{R}).$$

The Pauli exclusion principle requires this wave function to be antisymmetrical with respect to the exchange of the two nuclei, which implies that

$$\Psi_{\sigma_1\sigma_2}(\boldsymbol{R}) = \Psi_{\sigma_2\sigma_1}(-\boldsymbol{R}). \tag{4.231}$$

Since the interaction between the nuclear spins and the other degrees of freedom is weak, one can choose a base in which the orbital degrees of freedom and those relating to spin are factorized, and one obtains

$$\Psi_{\sigma_1\sigma_2}(\boldsymbol{R}) = \psi(\boldsymbol{R})\, \chi_{\sigma_1\sigma_2}. \tag{4.232}$$

One can moreover choose the orbital wave function $\psi(\boldsymbol{R})$ so that it is an eigenfunction of the angular momentum:

$$\psi(\mathbf{R}) = \psi^{(\mathrm{r})}(|\boldsymbol{R}|)\, Y_{\ell m}(\theta, \phi), \tag{4.233}$$

where θ and ϕ are the angular coordinates of $\boldsymbol{R}$. Let us recall that as a consequence of the transformation $\boldsymbol{R} \longrightarrow -\boldsymbol{R}$, one has $\theta \longrightarrow \pi - \theta$ and $\phi \longrightarrow \phi + \pi$, and therefore

$$Y_{\ell m}(\theta, \phi) \longrightarrow (-1)^{\ell} Y_{\ell m}. \tag{4.234}$$

Morever, the spin wave function can correspond to a total state of spin $s = 1$, and therefore be symmetrical, or to a spin $s = 0$, and be antisymmetrical. In order for the total wave function to be antisymmetrical, we must have

$$\ell = \begin{cases} 0, 2, 4, \ldots, & \text{if } s = 0; \\ 1, 3, 5, \ldots, & \text{if } s = 1. \end{cases} \tag{4.235}$$

On the other hand, the multiplicity of states with $s = 0$ is equal to 1, while that of states with $s = 1$ is equal to 3. The contribution of the rotational degrees of freedom to the specific heat of hydrogen can therefore be obtained by considering the rotational partition function:

$$\begin{aligned} \zeta_{\mathrm{rot}}(T) = &\sum_{\ell \text{ even}} (2\ell + 1) \exp\left[-\frac{\hbar^2}{2\mathcal{I}} \ell(\ell + 1) \right] \\ &+ 3 \sum_{\ell \text{ odd}} (2\ell + 1) \exp\left[-\frac{\hbar^2}{2\mathcal{I}} \ell(\ell + 1) \right] = \zeta_{\mathrm{even}}(T) + \zeta_{\mathrm{odd}}(T). \end{aligned} \tag{4.236}$$

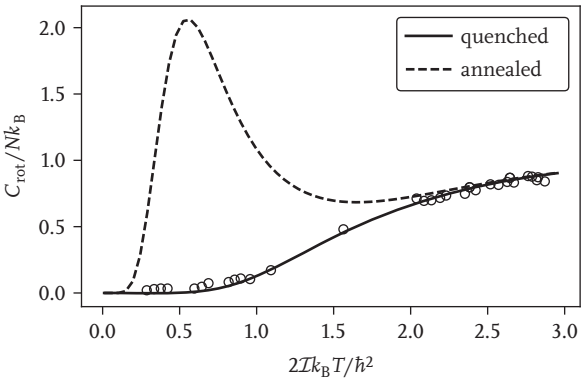


FIGURE 4.12. Rotational specific heat per particle (in units of k_B) of H_2. Circles represent experimental data. The dashed curve is the prediction of eq. (4.237), where one assumes that ortho- and para-hydrogen equilibrate at all temperatures. The full curve corresponds to eq. (4.238), which assumes that the ratio of ortho- to para-hydrogen remains quenched to its high-temperature value of 3:1. Based on [Wann66, pp. 219, 221].

From this expression, we obtain the rotational specific heat of hydrogen:

$$C_{\text{rot}} = Nk_B \frac{\partial}{\partial T}\left(T^2 \frac{\partial \ln \zeta_{\text{rot}}}{\partial T}\right). \tag{4.237}$$

This expression, however, does not agree with experimental data at low temperatures. In figure 4.12 the result of this calculation corresponds to the dashed line, while the circles represent experimental data.

This discrepancy is due to the fact that, since the interactions with the nuclear spins are very weak, states of different total nuclear spin do not equilibrate in the characteristic times of these specific-heat experiments. The quantities of molecules N_0 in the state $s=0$ and N_1 in the state $s=1$ must be considered as *independent thermodynamic variables*. It is as if molecular hydrogen were a mixture of two types of particles, **para-hydrogen**, with a nuclear spin equal to 0, and **ortho-hydrogen**, with a nuclear spin equal to 1.

In this fashion, one obtains the following expression for rotational specific heat:

$$C_{\text{rot}} = Nk_B T\left[\frac{N_0}{N}\frac{\partial^2(T\ln\zeta_{\text{even}})}{\partial T^2} + \frac{N_1}{N}\frac{\partial^2(T\ln\zeta_{\text{odd}})}{\partial T^2}\right], \tag{4.238}$$

where $\zeta_{\text{even}}(T)$ and $\zeta_{\text{odd}}(T)$ are defined in eq. (4.236).

This expression can be interpreted by saying that it is necessary to average the *logarithm* of the partition function over the "quenched" variables N_0 and N_1. The experimental data shown in figure 4.12 match fairly well the values

$$N_0 = \frac{N}{4}; \qquad N_1 = \frac{3N}{4}. \tag{4.239}$$

These are the equilibrium values at high temperatures. The prediction of eq. (4.238) with these values of N_0 and N_1 corresponds to the full line in figure 4.12, and matches well with the data collected in several experiments in the early twentieth century.

4.12 Chemical Equilibria in Gases

Let us consider a chemical reaction defined by the formula (2.162) of sec. 2.20, which we reproduce here:

$$\nu_1 \mathrm{X}_1 + \cdots + \nu_k \mathrm{X}_k \leftrightharpoons \nu_{k+1} \mathrm{X}_{k+1} + \cdots + \nu_r \mathrm{X}_r. \tag{4.240}$$

This expression can be rewritten in a more symmetrical manner:

$$\sum_i \nu_i \mathrm{X}_i = 0, \tag{4.241}$$

where the X_i denote the different chemical species, and the stoichiometric coefficients relative to the left-hand side of the reaction (the *reactants*) are negative. Thus the system can move along a line in the space identified by the number N_i of particles of chemical species i, given by the condition

$$\sum_i \nu_i \Delta N_i = 0. \tag{4.242}$$

We have seen in sec. 2.20 that the equilibrium condition with respect to these possible changes in particle numbers is expressed by

$$\sum_i \mu_i \nu_i = 0. \tag{4.243}$$

We can now use a relation that holds for gases at low density:

$$\mu = k_\mathrm{B} T \left[\ln \left[\rho \lambda_\mathrm{B}^3(T) \right] + \ln \zeta_\mathrm{int}(T) \right]. \tag{4.244}$$

Substituting this expression of μ in the equilibrium condition, we obtain the law of **mass action**:

$$\prod_i \rho_i^{\nu_i} = \prod_i \left[\lambda_{\mathrm{B},i}^3(T)\, \zeta_\mathrm{int}(T) \right]^{\nu_i} = K(T). \tag{4.245}$$

The quantity on the right-hand-side depends only on temperature and can in principle be calculated.

Recommended Reading

The arguments discussed in the present chapter can be found in all standard textbooks on statistical mechanics, such as K. Huang, *Statistical Mechanics* (2nd ed.), New York: J. Wiley & Sons, 1987; R. K. Pathria and P. D. Beale, *Statistical Mechanics* (3rd ed.), Amsterdam: Elsevier, 2011; and M. Plischke and B. Birgerson, *Equilibrium Statistical Mechanics* (3rd ed.), Singapore; World Scientific, 2006. I also recommend two exercise books: H. Krivine and J. Treiner, *La Physique Statistique en Exercices*, Paris: Vuibert, 2008, and Y.-K. Lim, *Problems and Solutions in Statistical Mechanics*, Singapore: World Scientific, 1990.

5 Phase Transitions

The lobster demands to be cut while alive.

—Brillat-Savarin

5.1 Liquid-Gas Coexistence and Critical Point

I borrow from M. E. Fisher this vivid description of the critical point of carbon dioxide [Fish83, p. 6]:

> The first critical point to be discovered was in carbon dioxide. Suppose one examines a sealed tube containing CO_2 at an overall density of about 0.5 gm/cc, and hence a pressure of about 72 atm. At a temperature of about 29°C one sees a sharp meniscus separating liquid (below) from vapor (above). One can follow the behavior of the liquid and vapor densities if one has a few spheres of slightly different densities close to 0.48 gm/cc floating in the system. When the tube is heated up to about 30°C one finds a large change in the two densities since the lighter sphere floats up to the very top of the tube, i.e., up into the vapor, while the heaviest one sinks down to the bottom of the liquid. However, a sphere of about "neutral" density (in fact "critical density") remains floating "on" the meniscus. There is, indeed, still a sharp interface between the two fluids, but they have approached one another closely in density. Further slight heating to about 31°C brings on the striking phenomenon of critical opalescence. If the carbon dioxide, which is quite transparent in the visible region of the spectrum, is illuminated from the side, one observes a strong intensity of scattered light. This has a bluish tinge when viewed normal to the direction of illumination, but has a brownish-orange streaky appearance, like a sunset on a smoggy day, when viewed from the forward direction (i.e., with the opalescent fluid illuminated from behind). Finally, when the temperature is raised a further few tenths of a degree, the opalescence disappears and the fluid becomes completely clear again. Moreover, the meniscus separating "liquid" from "vapor" has completely vanished: no trace of it remains! All differences between the two phases have gone: indeed, only one, quite homogeneous, "fluid" phase remains above the critical temperature ($T_c = 31.04$°C).

We can describe this behavior in a pressure–volume diagram like the one shown in figure 5.1. This diagram is based on the (approximate) van der Waals equation of state, which we discuss in the next section. Above a characteristic temperature value T_c, called

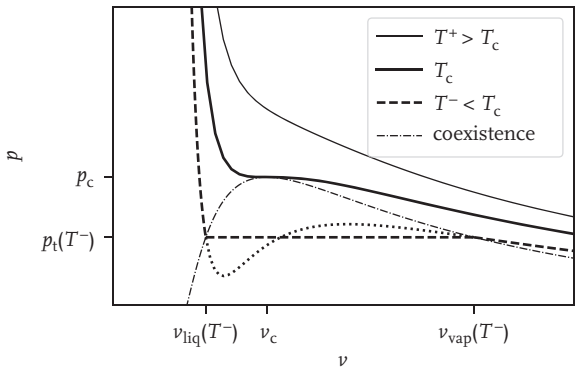


FIGURE 5.1. Pressure-volume isotherms of a fluid according to the van der Waals equation for three values of the temperature. The dotted line represents the non-physical part of the isotherm for $T^- < T_c$ in the coexistence region, and $p_t(T^-)$ is the transition pressure for $T = T^-$. The coexistence curve, $v_{liq}(p)$, and $v_{vap}(p)$ for $p < p_c$ are also shown.

the **critical temperature**, the pressure p is a monotonically decreasing function of the specific volume per particle $v = V/N$. For large values of v, p is inversely proportional to v, as prescribed by the equation of state for ideal gases. When p becomes very large, v cannot decrease beyond a certain value, which is dictated by the intrinsic volume of the molecule. As temperature decreases, the isotherm tends becomes flatter in the intermediate region, until for $T = T_c$ it shows a horizontal inflection point at the value $v = v_c = 1/\rho_c$. This density value is called the **critical density**, and the corresponding value p_c of the pressure is called the **critical pressure**. Below the critical temperature, the isotherm displays a flat region, between $v = v_{liq} = 1/\rho_{liq}$ and $v = v_{vap} = 1/\rho_{vap}$, where $\rho_{liq,vap}$ are the densities of the liquid and the vapor, respectively. Only one pressure value corresponds to all density values included between ρ_{vap} and ρ_{liq}: it is called the **transition pressure** (or **vapor pressure**), and we denote it by $p_t(T)$. What changes along this part of the isotherm is the fraction of liquid and vapor, respectively. The actual fractions at a given v are obtained by the **lever rule**:

$$x_{liq} = \frac{v_{vap} - v}{v_{vap} - v_{liq}}; \qquad x_{vap} = \frac{v - v_{liq}}{v_{vap} - v_{liq}}. \tag{5.1}$$

We thus obtain two curves, $v_{liq}(p_t)$ and $v_{vap}(p_t)$ that join at (v_c, p_c).

By plotting $p_t(T)$ vs. T in a temperature-pressure diagram, we obtain a line terminating at (T_c, p_c), as shown in fig. 5.2. Let us note that both axes in this diagram correspond to intensive quantities. The coexistence curve $p = p_t(T)$ in the (T, p) plane delimits the region of vapor on the left from that of liquid on the right. This curve terminates at (T_c, p_c), without exhibiting any singularities. We have thus obtained the **phase diagram** for the substance under study.

Other curves in the same plane describe the coexistence of fluid and solid phases—they meet in the triple point (T_{tr}, p_{tr}), where all three phases coexist. One should note that neither the solid-liquid nor the solid-vapor coexistence exhibits a critical point, unlike the liquid-vapor coexistence.

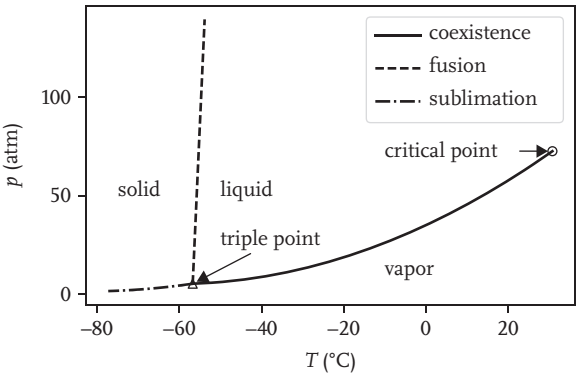


FIGURE 5.2. Phase diagram of CO_2 on the plane (T, p). Notice the liquid-vapor coexistence curve, which terminates at the critical point ($T_c = 30.98°C$, $p_c = 72.79$ atm), and the fusion (solid-liquid coexistence) and sublimation (solid-vapor coexistence) curves, which meet the preceding curve at the triple point ($T_{tr} = -56.57°C$, $p_{tr} = 5.11$ atm).

5.2 Van der Waals Equation

A first interpretation of this curious thermodynamic behavior was given in van der Waals' doctoral thesis [Waal73]. James C. Maxwell was so enthusiastic about this work (to which he made a substantial contribution [Maxw75], as we will see shortly) that he wrote that "it has certainly directed the attention of more than one enquirer to the study of the Low-Dutch language in which it is written." Van der Waals starts from the equation of state for ideal gases:

$$p = \frac{k_B T}{v}, \tag{5.2}$$

where $v = 1/\rho$ is the volume per molecule. He then takes two effects into account:

1. The effect of the intrinsic volume of the gas particles: the volume actually available to each particle is not equal to v, but is equal to $(v - b)$, where $b > 0$ (the **covolume**) is a parameter related to the intrinsic volume of the particles.

2. The effect of the reciprocal attraction among the particles: the pressure exerted by the particles on the walls is smaller than the actual pressure acting on them within the sample, because they tend to be attracted to the particles in the bulk of the fluid. In order to evaluate this effect, let us denote by $-u(r)$ the potential energy of a pair of molecules placed at a distance r apart. We then denote by u_b the total potential energy of a particle in the bulk of the fluid, due to the attraction by the other particles, and by u_w the corresponding value of the potential energy for a particle close to the wall. One then expects

   $$u_w = \frac{u_b}{2}, \tag{5.3}$$

   since only the particles within a half-space attract a particle close to the wall. Assuming that the density $\rho = 1/v$ is uniform, u_b is given by

   $$u_b = \frac{u_0}{v}, \tag{5.4}$$

where u_0 is the integral of the attractive potential $u(r)$:

$$u_0 = \int \mathrm{d}\boldsymbol{r}\, u(r) < 0. \tag{5.5}$$

This allows us to evaluate the local density near the wall, as a function of the bulk density $1/v$:

$$\rho_\mathrm{w} = \frac{1}{v}\, \exp\left(-\frac{u_\mathrm{w} - u_\mathrm{b}}{k_\mathrm{B} T}\right) = \frac{1}{v}\, \mathrm{e}^{u_0/2vk_\mathrm{B}T}. \tag{5.6}$$

Supposing that the pressure exerted by the molecules near the wall is given at least approximately by the ideal gas law, we obtain

$$\begin{aligned} p = \rho_\mathrm{w}\, k_\mathrm{B} T &= \frac{k_\mathrm{B} T}{v}\, \mathrm{e}^{u_0/2vk_\mathrm{B}T} \simeq \frac{k_\mathrm{B} T}{v}\left(1 + \frac{u_0}{2vk_\mathrm{B}T}\right) \\ &= \frac{k_\mathrm{B} T}{v} + \frac{u_0}{2v^2}. \end{aligned} \tag{5.7}$$

Putting the two effects together, we obtain the **van der Waals equation**:

$$\left(p + \frac{a}{v^2}\right)(v - b) = k_\mathrm{B} T, \tag{5.8}$$

where $a = -u_0/2 > 0$ and b are two parameters characteristic of the material.

The pressure predicted by this equation decreases monotonically with increasing v when the temperature is larger than a value T_c, given by

$$T_\mathrm{c} = \frac{8a}{27bk_\mathrm{B}}. \tag{5.9}$$

For $T = T_\mathrm{c}$, the curve exhibits a horizontal inflection at a value v_c of the volume per particle, given by

$$v_\mathrm{c} = 3b. \tag{5.10}$$

The corresponding value of the pressure is

$$p_\mathrm{c} = \frac{a}{27b^2}. \tag{5.11}$$

We can then identify T_c with the critical temperature, v_c with the critical volume per particle (inverse of the critical density ρ_c), and p_c with the critical pressure. Remarkably, if we introduce the **reduced variables**

$$T_\mathrm{r} = \frac{T}{T_\mathrm{c}}, \qquad v_\mathrm{r} = \frac{v}{v_\mathrm{c}}, \qquad p_\mathrm{r} = \frac{p}{p_\mathrm{c}}, \tag{5.12}$$

we obtain an equation in which all material-dependent parameters have disappeared:

$$\left(p_\mathrm{r} + \frac{3}{v_\mathrm{r}^2}\right)(3v_\mathrm{r} - 1) = 8T_\mathrm{r}. \tag{5.13}$$

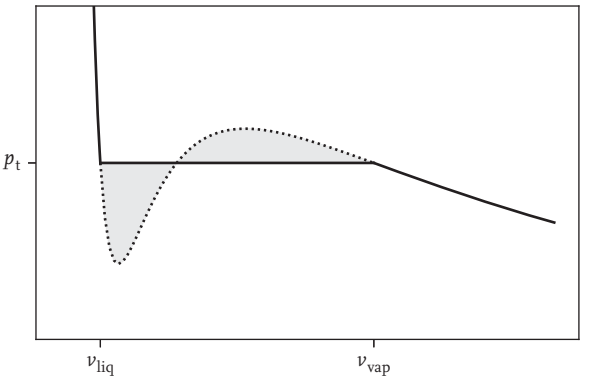


FIGURE 5.3. Maxwell construction. The areas delimited by the isotherm above and below the line $p = p_t = \text{const.}$ are equal.

Thus, as long as actual fluids obey, at least approximately, van der Waals's equation, the behavior of one substance can be mapped on that of another substance by a suitable rescaling. This is known as the **law of corresponding states**. It yields good results if one is not too close to the critical temperature.

For $T < T_c$ there is an interval of values of v that contains v_c, where $p(v, T)$ depends non-monotonically on v. In this case, for an interval of values of p, the curve intersects a horizontal line at three points. There are therefore, in principle, three values of v for the given values of temperature and pressure. However, the intermediate value of v is nonphysical, since the compressibility there is negative: if there is, for example, a slight fluctuation that increases v, the pressure increases, which causes v to increase further, thus moving the system further and further away from the point. On the other hand, the other two intersections are stable and we may, tentatively, identify the smaller one with v_{liq} and the larger one with v_{vap}.

The pressure of liquid and vapor in coexistence (the **vapor pressure** p_t) is a function of the temperature only. So the problem boils down to the identification of the vapor pressure. To tackle this problem, Maxwell [Maxw75] imagined a Carnot-like machine performing a reversible isothermal cycle, by first following the line $p = p_t = \text{const.}$ from, say, v_{liq} to v_{vap}, and then tracing the non-monotonic isotherm back from v_{vap} to v_{liq}. Since the temperature is kept constant during the cycle, by Carnot's theorem no heat can have been transformed into work (and conversely no work into heat). But the net work performed in the cycle is given by the difference in the areas delimited by the isotherm and the horizontal line $p = p_t$, above and below the line. Therefore the value of p_t is determined as that in which these two areas are equal. This is known as Maxwell's **equal-area construction**, and is shown in fig. 5.3.

This ingenious reasoning can be criticized, since there is no way to follow experimentally the instable part of the van der Waals isotherm, where $dp/dv > 0$. However, it yields the correct result. Indeed, liquid and the vapor coexist when their chemical potentials are equal. Now, for a simple fluid we have the Gibbs-Duhem equation (2.129),

$$d\mu = v\,dp - s\,dT. \tag{5.14}$$

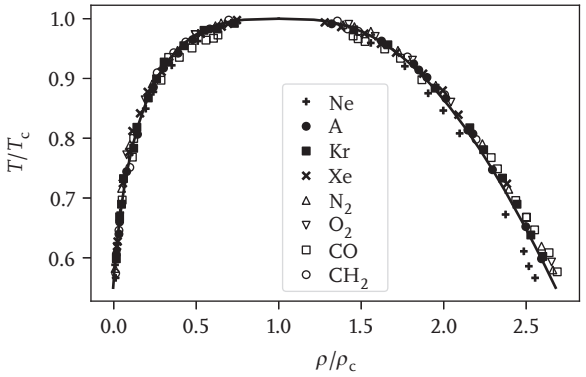


FIGURE 5.4. Coexistence curve, ρ/ρ_c, as a function of T/T_c for eight fluids in the vicinity of the critical temperature. The line corresponds to an empirical formula obtained by a fit of $\rho_{\mathrm{liq,vap}}/\rho_c = 1 + a(1 - T/T_c) \pm b(1 - T/T_c)^{1/3}$ to the data of argon. Adapted from ref. [Gugg45], with permission of AIP publishing.

Therefore the condition $\mu_{\mathrm{liq}} = \mu_{\mathrm{vap}}$ at $T = \text{const.}$ corresponds to the condition

$$\int_{v_{\mathrm{liq}}}^{v_{\mathrm{vap}}} \mathrm{d}p\, v = [pv]_{v_{\mathrm{liq}}}^{v_{\mathrm{vap}}} - \int_{v_{\mathrm{liq}}}^{v_{\mathrm{vap}}} \mathrm{d}v\, p = 0. \tag{5.15}$$

If we perform the integration along the van der Waals isotherm, this yields Maxwell's construction. However, since μ is a state function, we can evaluate the difference $\mu_{\mathrm{liq}} - \mu_{\mathrm{vap}}$ by integrating the Gibbs-Duhem equation along any path we wish, in particular along a path that goes around the critical point in the one-phase region, and that therefore does not encounter any unphysical states. Since we are given an analytic form of the equation of state, and we go back to the same value of the temperature, the result must be the same.

Exercise 5.1 Show that the Maxwell construction corresponds to the double tangent construction discussed in sec. 2.21.

More particularly, it follows from the van der Waals equation that for T close enough to T_c, the difference between the specific volumes of the vapor and the liquid vanishes as a power of $|T - T_c|$:

$$v_{\mathrm{vap}} - v_{\mathrm{liq}} \sim |T - T_c|^{\beta}, \qquad \beta = \frac{1}{2}. \tag{5.16}$$

Experimentally, this quantity actually behaves as a power of $|T - T_c|$, but with an exponent β closer to $1/3$ than to $1/2$. It is also interesting to note that the value of this exponent is apparently the same (within the errors) for several different liquids, and that the equation of state has an apparently universal form, as can be seen in figure 5.4.

The isothermal compressibility of a fluid following the van der Waals equation exhibits a singularity at the critical point, which can be verified experimentally:

$$\chi = -\frac{1}{v}\left.\frac{\partial v}{\partial p}\right)_T \sim |T - T_c|^{-\gamma}, \qquad \gamma = 1. \tag{5.17}$$

This holds both for $v = v_c$ above T_c and for $v = v_{\mathrm{liq,vap}}(T)$ below T_c. However, in actual fluids, the exponent γ is closer to 1.25.

To evaluate the specific heat of the van der Waals fluid, we exploit the fact that

$$\frac{\partial^2 S}{\partial T\,\partial V} = \frac{\partial}{\partial T}\left.\frac{\partial p}{\partial T}\right)_V , \tag{5.18}$$

where we have used a Maxwell relation. We obtain therefore

$$\left.\frac{\partial C_V}{\partial V}\right)_T = T\frac{\partial^2 S}{\partial V\,\partial T} = T\frac{\partial^2 S}{\partial T\,\partial V} = T\left.\frac{\partial^2 p}{\partial T^2}\right)_V . \tag{5.19}$$

Now, T appears linearly in the expression of $p(v, T)$ given by the van der Waals equation. Therefore the heat capacity of the van der Waals fluid at fixed volume is constant. We can evaluate it at high temperatures, where the gas follows the ideal gas law. Below T_c, we must take into account that the transition pressure depends on the temperature, and the internal energy depends on $v_{\mathrm{liq,vap}}(T)$. We then obtain a discontinuity in C_V in approaching T_c with $v = v_c$ from below and from above. Again, the behavior of actual fluids is different, since one observes a weak power law singularity of the form

$$C_V \sim |T - T_c|^{-\alpha} , \tag{5.20}$$

where the exponent α is the same (within the errors) for all fluids and lies between 1/8 and 1/9.

Exercise 5.2 Evaluate the discontinuity of the specific heat in a fluid following the van der Waals equation for $v = v_c$ and T close to T_c:

1. Express the specific heat per particle at constant volume, evaluated at the critical volume v_c, in terms of the fractions x_{liq} and x_{vap} at coexistence, of the specific heats for the coexisting liquid and vapor phases, of the derivatives $\partial p_{\mathrm{liq}} = \partial v_{\mathrm{liq}})_T$ and $\partial p_{\mathrm{vap}} = \partial v_{\mathrm{vap}})T$ of the equation of state at coexistence, and of the derivatives $\mathrm{d}v_{\mathrm{liq}}/\mathrm{d}T)_{\mathrm{coex}}$ and $\mathrm{d}v_{\mathrm{vap}}/\mathrm{d}T)_{\mathrm{coex}}$ of the specific volumes of the two phases at coexistence.

2. Show that the specific heat at constant volume of a van der Waals fluid in a homogeneous phase does not depend on the temperature.

3. Evaluate the relevant quantities mentioned above close to the critical point in a van der Waals fluid. Deduce the jump in the specific heat per particle at constant volume equal to the critical volume around the critical temperature.

It is convenient to use *reduced units*, namely, to measure quantities such as the temperature T, the specific volume v, and the pressure p in units of the corresponding critical values T_c, v_c, and p_c.

5.3 Binary Mixtures

Another system that exhibits analogous behavior is a fluid mixture composed of two chemical species. One usually uses organic compounds, such as aniline-cyclohexane, often

chosen because its critical point is close to the ambient temperature at atmospheric pressure. We refer to one of the components as A and to the other as B. By fixing the pressure, it is possible to observe the coexistence of *three* phases within a certain temperature interval, and they are arrayed from the bottom up in a decreasing order of density: an α phase rich in A, a β phase rich in B, and a vapor phase. As the temperature varies, the compositions (and therefore also the densities) of the α and β phases move closer to each other until, for a particular temperature value T_c (as long as we have selected the correct composition of the sample) the two liquid phases are replaced by a homogeneous phase. The point at which this occurs is called the **consolution point**.

The behavior of this system is fairly similar to that of the liquid–gas system. More particularly, it too exhibits critical opalescence in the proximity of the consolution point. Let us note some differences:

1. The quantity that distinguishes the two phases is the relative concentration of one of the components—A, for example. Let us denote the molar fraction of A by f_A and its value in phase α by $f_A^{(\alpha)}$. Then, $f_A^{(\alpha)} - f_A^{(\beta)}$ approaches 0 as T approaches T_c, from below, when the overall fraction of A is equal to its value at the consolution point.

2. One can observe, however, that the exponent β that describes the behavior of this system is also close to 1/3 and indistinguishable from the value it assumes in simple fluids.

3. Fairly frequently, the consolution point corresponds to a temperature *lower* than the coexistence temperature. In other words, a homogeneous phase is replaced by a phase coexistence when the sample is heated. This does not occur with simple fluids.

5.4 Lattice Gas

In order to obtain a more microscopic description of phase coexistence, let us build a simple statistical model of a "fluid." We consider a system of N particles, described by the Hamiltonian

$$\mathcal{H}(\boldsymbol{p}^N, \boldsymbol{r}^N) = K(\boldsymbol{p}^N) + U(\boldsymbol{r}^N), \tag{5.21}$$

where $\boldsymbol{p}^N = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N)$, $\boldsymbol{r}^N = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$, and $K(\boldsymbol{p}^N)$ is the kinetic energy

$$K(\boldsymbol{p}^N) = \sum_{i=1}^{N} \frac{p_i^2}{2m}, \tag{5.22}$$

while $U(\boldsymbol{r}^N)$ is the potential energy, which depends only on the positions of the particles. We then calculate the partition function, obtained by generalizing the expression (3.96) to a system of interacting particles:

$$Z_N = \frac{1}{N!} \int \prod_{i=1}^{N} \left(\frac{\mathrm{d}\boldsymbol{p}_i \mathrm{d}\boldsymbol{r}_i}{h^3} \right) \exp\left(-\frac{K(\boldsymbol{p}^N) + U(\boldsymbol{r}^N)}{k_B T} \right). \tag{5.23}$$

Here, h is Planck's constant, which appears to account for the volume of phase space occupied by the quantum states, and the factor $1/N!$ takes into account the indistiguishability

of the particles. The integral over the momenta $\boldsymbol{p}_i$ factorizes and results in λ_B^{-3N}, where λ_B is the thermal de Broglie wavelength. In this fashion, the partition function is expressed in terms of an integral over the positions $\boldsymbol{r}^N$ of the particles:

$$Z_N = \frac{\lambda_B^{-3N}}{N!} \int \prod_{i=1}^{N} \mathrm{d}\boldsymbol{r}_i \, \exp\left(-\frac{U(\boldsymbol{r}^N)}{k_B T}\right) = \frac{\lambda_B^{-3N}}{N!} \, Q_N, \tag{5.24}$$

where we have defined the **configuration integral**;

$$Q_N = \int \mathrm{d}\boldsymbol{r}^N \, \exp\left(-\frac{U(\boldsymbol{r}^N)}{k_B T}\right). \tag{5.25}$$

The interaction between particles is usually described by a potential that exhibits a hard core—it becomes large and positive when the particles get closer than a certain distance a_0, where $a_0/2$ is the particle radius—and an attractive tail, that is, it becomes negative at intermediate distances, eventually vanishing when the particles are far apart. We can provide a drastically simplified description of this state of affairs with the following hypotheses:

1. The particles can occupy only the sites on a simple cubic lattice of $\mathcal{N}$ sites with lattice constant equal to a_0, where $\mathcal{N}$ is given by

$$\mathcal{N} = \frac{V}{a_0^3}. \tag{5.26}$$

2. At every site i placed at $\boldsymbol{r}_i$, there can be one particle at most.

3. If two particles occupy nearest-neighbor lattice sites, that is, sites i and j such that $|\boldsymbol{r}_i - \boldsymbol{r}_j| = a_0$, the potential energy of the system is reduced by a quantity ϵ.

Let us denote the state of occupation of site i by a variable τ_i, which is equal to 0 if the position is empty and to 1 if it is occupied. Then the partition function assumes the form

$$Z_N = \lambda_B^{-3N} \sum_{\tau} \delta_{N,\sum_i \tau_i} \exp\left(\frac{\epsilon}{k_B T} \sum_{<i,j>} \tau_i \tau_j\right), \tag{5.27}$$

where $\tau = (\tau_1, \ldots, \tau_{\mathcal{N}})$ is the collection of the values of τ for each point i of the lattice and the expression $\sum_{<i,j>}$ indicates that the sum runs over all pairs of nearest-neighbor sites on the lattice. The combinatorial factor $N!$ has disappeared because this expression is invariant with respect to the permutation of particle labels and the Kronecker delta imposes the condition that the total number of particles in the system is exactly N:

$$\sum_i \tau_i = N. \tag{5.28}$$

The model defined in this manner is called the **lattice gas**.

In order to make this expression more manageable, it is useful to eliminate the Kronecker delta—one can do this by passing to the grand canonical ensemble. In this fashion, we obtain the grand partition function

$$Z_{GC} = \sum_{N=0}^{\infty} \exp\left(\frac{\mu N}{k_B T}\right) \, Z_N. \tag{5.29}$$

The grand partition function Z is related to the pressure p by the equation

$$\ln Z_{\mathrm{GC}} = \frac{pV}{k_{\mathrm{B}}T}, \tag{5.30}$$

where V is the volume of the system. It is easy to see that Z can be expressed as follows:

$$Z_{\mathrm{GC}} = \sum_{\tau} \exp\left(-\frac{\mathcal{H}(\tau)}{k_{\mathrm{B}}T}\right), \tag{5.31}$$

where the sum runs over all possible configurations τ of the system and the Hamiltonian $\mathcal{H}(\tau)$ is defined by

$$\mathcal{H}(\tau) = -\sum_{<ij>} \epsilon \tau_i \tau_j - \sum_i \bar{\mu} \tau_i. \tag{5.32}$$

In this equation, we introduced the expression

$$\bar{\mu} = \mu - 3k_{\mathrm{B}}T \ln \lambda_{\mathrm{B}}, \tag{5.33}$$

to take into account the effect of the integral over momenta.

The particle density per site is given by the quantity

$$\rho = \frac{1}{\mathcal{N}} \sum_i \langle \tau_i \rangle . \tag{5.34}$$

For T smaller than a critical temperature T_{c}, ρ exhibits a discontinuity for $\bar{\mu} = \bar{\mu}_{\mathrm{t}}(T)$:

$$\rho^+ = \lim_{\bar{\mu} \to \bar{\mu}^+} \rho > \rho^- = \rho^+ = \lim_{\bar{\mu} \to \bar{\mu}^-} \rho. \tag{5.35}$$

The discontinuity tends to zero when the temperature gets close to the critical one. If the model correctly describes the experimental data, we expect the discontinuity to tend to zero as a power of $|T - T_{\mathrm{c}}|$:

$$\rho^+ - \rho^- \sim |T - T_{\mathrm{c}}|^{\beta} . \tag{5.36}$$

5.5 The Ising Model

The model we just described cannot be solved exactly. It is possible, however, to manipulate it to highlight some interesting properties. Consider a lattice gas model with $\mathcal{N}$ sites. Let us change the variables describing the microstate by the mapping $\tau \longrightarrow \sigma$, defined by

$$\sigma_i = 2\tau_i - 1, \qquad \forall i. \tag{5.37}$$

This variable assumes the values -1 (when $\tau_i = 0$) and $+1$ (when $\tau_i = 1$). Expressing the τ_i as a function of the σ_i, and substituting in the Hamiltonian (5.32), we obtain (up to an irrelevant additive constant) the expression

$$\mathcal{H}(\sigma) = -\sum_{<ij>} J\sigma_i\sigma_j - \sum_i h\sigma_i. \tag{5.38}$$

Here we have defined

$$J = \epsilon/4, \tag{5.39}$$

$$h = (\bar{\mu} + \epsilon\zeta)/2, \tag{5.40}$$

where ζ is the **coordination number** of the lattice, that is, the number of nearest neighbors of a given lattice site. In the case of the simple cubic lattice, $\zeta = 6$ (and it is equal to $2d$ in the case of a cubic lattice in d dimensions). This Hamiltonian can also describe binary mixtures, if one stipulates that $\sigma_i = +1$ corresponds to the presence of a molecule of type A in i, $\sigma_i = -1$ to a molecule of type B, and $2h = \mu_B - \mu_A$.

This expression makes explicit the fact that $\mathcal{H}(\sigma)$, for $h = 0$, is invariant with respect to the transformation $\sigma \longrightarrow \mathsf{I}\,\sigma$, where

$$\mathsf{I}\,\sigma_i = -\sigma_i, \qquad \forall i. \tag{5.41}$$

We can express this fact by saying that when $h = 0$, the Hamiltonian $\mathcal{H}(\sigma)$ is **symmetrical** with respect to the transformation group G composed of I and the identity E. Since $\mathsf{I}^2 = \mathsf{E}$, it is not necessary in this case to consider other transformations.

The model we obtained is the most famous model of statistical mechanics, the **Ising model**. It was invented by W. Lenz and proposed by him as a thesis to his student E. Ising [Isin25], who solved it in one dimension. It was actually designed to describe the behavior of ferromagnetic systems, that is, the fact that they exhibit a nonvanishing magnetic moment at temperatures below a given temperature T_c. In this case, σ_i is a constant times the z component of the magnetic moment of an electron localized at site i, and h is proportional to the applied magnetic field. The factor J describes the tendency of electrons in nearest-neighbor atoms to align themselves in parallel because of the exchange interaction—it is therefore known as the **exchange integral**. In spite of the quantum origin of the magnetic moment and of the interaction, the Ising model is treated as a classical model, since all the variables that appear in it commute.

5.6 Lee-Yang Theory of Phase Transitions

In two papers of 1952, C. N. Yang and T. D. Lee provide a mathematically rigorous approach to the condensation phenomenon, and apply their method to the Ising model [Yang52, Lee52]. They describe a simple fluid as a system of interacting particles, and assume that the interaction potential between particles is made of a repulsive hard core and an attractive tail of finite range. As a consequence of the hard core, the number N of particles that can be accommodated in a region of finite volume V is bounded. Therefore, the grand partition function Z is a polynomial of finite degree $\mathcal{N}$ in the fugacity $z = e^{\mu/k_B T}$:

$$Z_{GC}(z, V) = \sum_{N=0}^{\mathcal{N}} z^N Z_N, \tag{5.42}$$

where Z_N is the canonical partition function of a system with N particles:

$$Z_N = \frac{1}{N!\lambda_B^{dN}} \int \prod_{i=1}^{N} d\boldsymbol{r}_i \, e^{-U(\boldsymbol{r}^N)/k_B T}. \tag{5.43}$$

Here d is the dimensionality of the space, λ_B is the de Broglie thermal wavelength, and $U(\boldsymbol{r}^N)$ is the potential energy of the configuration $\boldsymbol{r}^N = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$:

$$U(\boldsymbol{r}^N) = \sum_i u_0(\boldsymbol{r}_i) + \frac{1}{2}\sum_{i\neq j} u(|\boldsymbol{r}_i - \boldsymbol{r}_j|). \tag{5.44}$$

Since Z_{GC} is a polynomial in z of degree equal to $\mathcal{N}$, it has $\mathcal{N}$ zeros (counted with their multiplicities). On the other hand, since

$$Z_N \geq 0, \quad \forall N, \tag{5.45}$$

there are no zeros of Z_{GC} on the positive real z axis, that is, for physical values of the fugacity z. Therefore $\ln Z_{GC}$ is analytic in a region of the complex z plane that includes the entire positive real axis. Now, the equation of state $p = p(\rho, T)$, where ρ is the particle density, is parametrically given by

$$\begin{aligned} \frac{p}{k_B T} &= \frac{1}{V}\ln Z_{GC}(z, V); \\ \rho &= \frac{1}{V} z \frac{\partial}{\partial z}\ln Z_{GC}(z, V). \end{aligned} \tag{5.46}$$

We thus obtain $p > 0$ and

$$\frac{\partial p}{\partial \rho} = \frac{\partial p}{\partial z}\frac{\partial z}{\partial \rho} = \frac{k_B T}{V}\frac{z}{\langle \rho^2\rangle - \langle \rho\rangle^2} \geq 0. \tag{5.47}$$

Therefore there are no singularities in the equation of state for the physical values of z, and this begs the question of how these singularities could arise. Yang and Lee's theory explains the mechanism producing these singularities in the thermodynamic limit, which in this case amounts to $V \to \infty$, $z = \text{const}$. Indeed, as the volume V increases, also the number $\mathcal{N}$ of zeros increases, and some of these zeros become closer and closer to the real axis, eventually pinching it.

Yang and Lee's results can be summarized in two theorems:

Theorem 1: The limit

$$\Omega(z) = \lim_{V\to\infty} \frac{1}{V}\ln Z_{GC}(z, V) \tag{5.48}$$

exists for all $z > 0$, and is a continuous, nondecreasing function of z. This limit is also independent of the shape of the container, as long as its surface area does not increase faster than $V^{(d-1)/d}$.

Theorem 2: Let a region R of the complex z plane contain a segment of the positive real z axis, but no zeros of $Z_{GC}(z, V)$ for any V. Then in this region the quantity $\ln Z_{GC}/V$ converges *uniformly* to its limit as $V \to \infty$, and the limit is analytic for all z in R.

I am not going to provide the proofs of these theorems; they can be found in the original papers. As a consequence of these theorems, the transition is determined by the behavior of the zeros of the partition function $Z_{GC}(z, V)$. Indeed, since the convergence to the limit is uniform in each zero-free region R, we can interchange the order of the limits in the

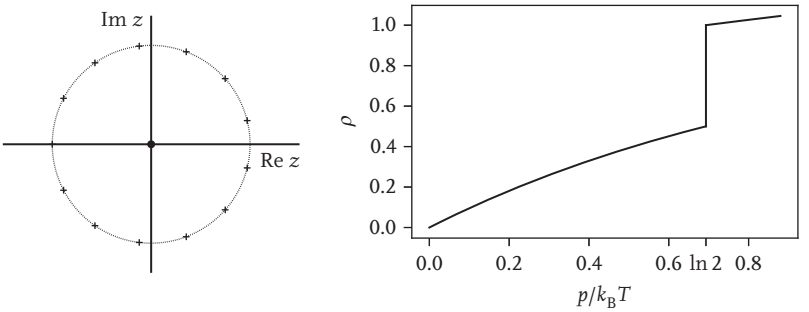


FIGURE 5.5. Left: Example of the distribution of zeros of the grand partition function $Z_{GC}(z, V) = (1+z)^V(1+z^{\alpha V})$ for $\alpha = 1/2$ and finite volume $V = 26$. The dotted line represents the unit circle. Right: Equation of state ρ vs. $p/k_B T$ for the same system, exhibiting a discontinuity for $z = 1$.

equation for ν and obtain the parametric equation of state in the form

$$\frac{p(z)}{k_B T} = \Omega(z); \qquad \rho = z\frac{\partial}{\partial z}\Omega(z). \tag{5.49}$$

The properties $p(z) > 0$ and $\partial p/\partial\rho > 0$ also hold in the thermodynamic limit.

A phase transition occurs when the zeros of $Z_{GC}(z, V)$ approach the real positive z axis for $V \to \infty$. Assume that no zeros approach the real axis for $z < z_0$ or for $z > z_0$, but that there are zeros with $\mathrm{Re}\, z \simeq z_0$ and $\mathrm{Im}\, z \sim V^{-1}$, which pinch the real axis at z_0. We then have two zero-free regions: R_1, given by $0 \le z < z_0$, and R_2, with $z_0 < z < \infty$. Then $p(z)$ is analytic in both R_1 and R_2 and continuous at z_0, and $\rho(z)$ is analytic in both regions, but may be discontinuous at z_0. Since $\rho(z)$ is an increasing function of z, it may make an upward jump when z crosses z_0. We thus obtain a discontinous phase transition. In this case, since the jump is the limit of continuous curves, all values of ρ between the extrema of the jump can be realized.

As an example, consider a hypothetical system whose grand partition function is given by $Z_{GC}(z, V) = (1+z)^V(1+z^{\alpha V})$, with $\alpha = 1/2$. The zeros of Z_{GC} for $V = 26$ are shown in figs. 5.5, left, while the right panel shows the corresponding equation of state. In this case, the zeros accumulate on the unit circle, leaving a gap of width of order V^{-1} at $z = z_0 = 1$. Thus we have R_1 inside the unit circle, and R_2 outside it. The transition pressure is given by $p = k_B T \ln 2$.

Exercise 5.3 (Phase coexistence) Consider the grand canonical partition function $Z_{GC}(z, V)$ given above, where α is a positive parameter.

1. Evaluate the density $\rho = N/V$ and show that the system exhibits phase coexistence. Evaluate the two densities at coexistence.

2. Evaluate the equation of state in the low-density phase ("gas") and the pressure at phase coexistence.

Things can be discussed more precisely in the lattice-gas model of $\mathcal{N}$ sites. Assume that no more than one particle can occupy a single lattice site, and that the interaction

potential is negative and of finite range, that is, it is not necessarily nearest-neighbor. Then the model can be mapped into an Ising model by transforming the occupation number τ_i of lattice site i into the spin variable σ_i by

$$\sigma_i = 2\tau_i - 1. \tag{5.50}$$

With this transformation, the Hamiltonian of the model takes the form

$$\mathcal{H}(\sigma) = -\frac{1}{2}\sum_{i\neq j} J_{ij}\sigma_i\sigma_j + h\sum_i \sigma_i, \tag{5.51}$$

where $J_{ij} = J_{ji} \geq 0$, and vanishes if i and j are sufficiently far from each other, and h is a linear function of the chemical potential μ. Defining the fugacity z by

$$z = \mathrm{e}^{h/k_\mathrm{B}T}, \tag{5.52}$$

we see that the partition function $Z(z, \mathcal{N})$ exhibits the symmetry

$$Z(z, \mathcal{N}) = Z(1/z, \mathcal{N}). \tag{5.53}$$

Since the zeros of Z appear in complex-conjugate pairs, this implies that all zeros must lie on the unit circle. If we denote the arguments these zeros by θ_k, we have

$$Z(z, \mathcal{N}) = \mathcal{Z}\prod_k \left(z - \mathrm{e}^{\mathrm{i}\theta_k}\right), \tag{5.54}$$

where

$$\mathcal{Z} = \left[\prod_k (-\mathrm{e}^{\mathrm{i}\theta_k})\right]^{-1}. \tag{5.55}$$

In the thermodynamic limit, the roots become continuously distributed on the unit circle. We introduce the root density $g(\theta)$ such that

$$g(\theta)\,\mathrm{d}\theta = \lim_{\mathcal{N}\to\infty}\frac{1}{\mathcal{N}}\int_\theta^{\theta+\mathrm{d}\theta}\mathrm{d}\theta' \sum_k \delta(\theta_k - \theta'). \tag{5.56}$$

Symmetry implies that $g(-\theta) = g(\theta)$. We can then express the equation of state via

$$\begin{aligned}\frac{p}{k_\mathrm{B}T} &= \int_0^\pi \mathrm{d}\theta\, g(\theta)\, \ln\left(1 - 2z\cos\theta + z^2\right);\\ \rho &= 2z\int_0^\pi \mathrm{d}\theta\, g(\theta)\,\frac{z - \cos\theta}{1 - 2z\cos\theta + z^2}.\end{aligned} \tag{5.57}$$

We see from this expression that the only point where singularities can occur is $z = 1$, and that this only happens if $g(0) > 0$. Therefore the transition can only take place at $z = 1$ and, if this happens, there will be different functions for $z > 1$ and $z < 1$, since the function in one region cannot be analytically continued in the other region.

Exercise 5.4 An electrostatic analog of the equation of state (5.57) of the Lee-Yang model can be found as follows:

1. Consider a circular cylinder of unit radius perpendicular to the complex z plane, cutting it on the unit circle. Suppose that the cylinder is charged with a surface charge density $g(\theta)$ that depends only on the angle θ and that satisfies $g(-\theta) = g(\theta)$. Let then $\Phi(z)$ and $E(z)$ be respectively the electrostatic potential and the electric field at a point z on the real axis. Show that we then have

$$\frac{p}{k_B T} = -\frac{1}{2}\Phi(z); \qquad \rho = \frac{1}{2} z E(z).$$

2. Assume $g(0) \neq 0$. Show by electrostatic arguments that $p(z)$ is continuous at $z = 1$, but ρ jumps discontinuously. By means of Gauss's theorem, show that the discontinuity is given by $\Delta\rho = 2\pi g(0)$.

5.7 Symmetry Breaking

We can now look at phase coexistence from another point of view. In the case of the ferromagnet, the two coexisting phases are equivalent from a physical standpoint and differ only because of the prevailing direction of their spin alignment. In other words, although the Hamiltonian is invariant with respect to the transformation group G, the thermodynamic state is not. This situation is called **spontaneous symmetry breaking**. It can be manifested by identifying some thermodynamic quantities that should vanish in a thermodynamic state that is symmetrical with respect to the transformation group G, but in fact do not vanish.

Let us consider, for instance, the **magnetization** M in the Ising model of N sites, defined by

$$\langle M \rangle = \sum_i \langle \sigma_i \rangle = \frac{1}{Z} \sum_\sigma \left(\sum_i \sigma_i \right) \exp\left[-\frac{\mathcal{H}(\sigma)}{k_B T} \right], \tag{5.58}$$

where the first sum runs over all 2^N configurations $\sigma = (\sigma_1, \ldots, \sigma_N)$ of the system, $\mathcal{H}(\sigma)$ is the Ising Hamiltonian (5.38), and Z is the partition function ($Z = \sum_\sigma \exp[-\mathcal{H}(\sigma)/k_B T]$).

Let us assume that $h = 0$, and let us consider the contributions to $\langle \sigma_i \rangle$. For each configuration σ in which $\sigma_i = +1$, there is a configuration $\mathrm{I}\sigma = -\sigma$ in which $\sigma_i = -1$. Since $\mathcal{H}(\mathrm{I}\sigma) = \mathcal{H}(\sigma)$, this configuration will, in principle, have the same Boltzmann weight as the preceding one. We can therefore expect $\langle \sigma_i \rangle$ to always vanish. But then how can phase coexistence occur? The point is that in certain circumstances, a small nonsymmetrical perturbation is sufficient to cause a macroscopic violation of symmetry. It is clear that states in which the spin is prevalently oriented up or down have a much lower energy than those in which the spins are arranged in a disordered fashion. The latter are much more numerous, and therefore prevail at high temperatures. But at low temperatures, they are not accessible (they cost too much energy). Therefore the spins tend be aligned prevalently in one direction: which direction is indifferent from an energetic point of view. A small perturbation (due to boundary conditions that slightly favor the up direction, or due to a very small magnetic field, for instance) is sufficient to make the system prefer one direction over the other.

It is possible in this manner to obtain a thermodynamic state in which magnetization is not equal to zero, even though the external macroscopic field h vanishes:

$$\langle M\rangle = \sum_i \langle \sigma_i\rangle = Nm \neq 0. \tag{5.59}$$

The symmetry of the Hamiltonian implies that if $M_0 = \langle M\rangle \neq 0$ defines a state of thermodynamic equilibrium, there exists another state of thermodynamic equilibrium (which corresponds to the same values of the intensive variables, but perhaps to different boundary conditions) where the magnetization is equal to $-M_0$. Therefore, if $\langle M\rangle \neq 0$ for $h = 0$, more than one equilibrium state, characterized by different values of the magnetization, can coexist.

The value of the magnetization can be fixed by means of an artifice called **quasi-average**. One applies an external field h—say, a positive one—which *explicitly* breaks the symmetry, and one evaluates the magnetization in the limit in which this field goes to zero:

$$Nm_0(T) = \lim_{h\to 0^+} \langle M\rangle\,. \tag{5.60}$$

Because of the symmetry $\mathcal{H}(\sigma, h) = \mathcal{H}(\mathsf{I}\sigma, -h)$, one has

$$\lim_{h\to 0^-} \langle M\rangle = -Nm_0(T). \tag{5.61}$$

Therefore, $m_0(T)$ is a measure of the degree of spontaneous symmetry breaking. This quantity is known as the **spontaneous magnetization** per particle and vanishes at the critical point.

The formal equivalence between the Ising model and the lattice gas allows us to identify the coexistence curve for the latter in the plane $(\bar{\mu}, T)$:

$$\bar{\mu}_t(T) = -\epsilon\zeta\,. \tag{5.62}$$

5.8 The Order Parameter

Let us now consider the same problem from a slightly more abstract point of view. Let us suppose that we have a system described by the Hamiltonian $\mathcal{H}(\sigma)$, which is invariant with respect to a transformation group G applied to the microscopic states σ:

$$\mathcal{H}(\mathsf{g}\sigma) = \mathcal{H}(\sigma), \qquad \forall \mathsf{g}\in\mathsf{G}, \quad \forall\sigma\,. \tag{5.63}$$

If the thermodynamic state is not invariant with respect to all elements of the group, one says that spontaneous symmetry breaking is present. In order to define the degree of symmetry breaking, one considers an observable M that has the following properties:

1. M is extensive: its value for a system composed of several subsystems is equal to the sum of the values pertaining to each subsystem.

2. M transforms according to a nontrivial representation of G:

   $$M(\mathsf{g}\sigma) = \mathsf{T}(\mathsf{g})M(\sigma), \qquad \forall \mathsf{g}\in\mathsf{G}, \quad \forall\sigma, \tag{5.64}$$

   where $\mathsf{T}(\mathsf{g})$ is a linear operator belonging to a nontrivial representation of G.

Then a nonvanishing value of $\langle M \rangle$ signals spontaneous symmetry breaking. With the quasi-average method, by choosing the symmetry-breaking fields in different directions, it is possible to obtain different values of $\langle M \rangle$ that correspond to the same value of temperature (and of other possible intensive parameters that do not break the symmetry). Because of symmetry, these values can be obtained from one another by applying a $\mathsf{T}(\mathsf{g})$ transformation.

A quantity endowed with these characteristics is called an **order parameter**. The introduction of this concept into phase transition theory is due to Lev D. Landau (see [Land80, ch. XIV]).

Let us now consider a certain number of systems, the transitions that characterize them, and the corresponding order parameters.

1. **Ising antiferromagnet**: Consider an Ising model on a simple cubic lattice, but assume that the interaction constant J between neighboring spins is negative. In the lowest energy states, the neighboring spins tend to be aligned in antiparallel fashion. In this case, both the reflection symmetry ($\sigma_i \longrightarrow -\sigma_i$) and the translational symmetry ($\sigma_i \longrightarrow \sigma_{\tau(i)}$, such that $\boldsymbol{r}_{\tau(i)} - \boldsymbol{r}_i = \text{const.}, \forall i$) are spontaneously broken at the same time.

   One can then use the **staggered magnetization** as the order parameter. In our case, one can identify two intertwined sublattices, such that each point of each sublattice is surrounded by nearest-neighbor points belonging to the other sublattice. Let us now assign to each point i a factor ε_i, whose value is $+1$ in one sublattice and -1 in the other. The staggered magnetization is then defined by

$$N(\sigma) = \sum_i \varepsilon_i \sigma_i. \tag{5.65}$$

   The symmetry group G is still the group with two elements (E, I) that correspond to spin identity and spin inversion. One also has $N(\mathsf{t}\sigma) = -N(\sigma)$ if t is the mapping $\sigma_i \longrightarrow \sigma_{\tau(i)}$ defined above, satisfying the condition $\left|\boldsymbol{r}_{\tau(i)} - \boldsymbol{r}_i\right| = a_0$, that is, a translation shifting the whole lattice by one lattice constant in any coordinate direction.

2. **Vector order parameter**: The Ising model describes a ferromagnet in which the spins are constrained to align parallel or antiparallel to a fixed direction. One often finds situations where the spins have more degrees of freedom and can point in arbitrary directions in a plane or in space. (We disregard here the quantum nature of the spins.) For instance, for a spin free to point in any direction in space, if we denote its components by σ_i^α, we can define the order parameter $M(\sigma) = (M^\alpha(\sigma))$ as follows:

$$M^\alpha(\sigma) = \sum_i \sigma_i^\alpha, \qquad \alpha \in \{x, y.z\}. \tag{5.66}$$

   More generally, one can consider vectorial order parameters with n components, which transform under rotations in n dimensions, represented by elements of the group $\mathsf{O}(n)$. The generalization of the Ising model for $n = 2$ is often called the **XY model**, while that for $n = 3$ is called the **Heisenberg model**.

3. **Einstein condensation**: Sometimes the order parameter is not directly accessible to the experiment. In the Einstein condensation, for example, the order parameter is the condensate wave function, and the broken symmetry is the gauge symmetry of the first kind, which

expresses the invariance with respect to the multiplication of the wave functions by a phase:

$$\Psi \longrightarrow e^{i\alpha}\Psi. \tag{5.67}$$

The gauge transformation is isomorphic to a rotation in the complex plane, and therefore the order parameter can be identified as a vector order parameter with $n = 2$.

There are more complex situations (for example, in liquid crystals) that require the introduction of less intuitive order parameters. What needs to be emphasized is that the identification of the order parameter can be a fairly difficult problem by itself, because there is no general method that allows one to identify it a priori. One needs to be guided by physical intuition.

When the conditions for state coexistence are satisfied (for example, $T < T_c$ and $h = 0$), the equilibrium state is not unique in the thermodynamic limit. Then the space of the possible equilibrium thermodynamic states contains all convex combinations of equilibrium states, that is, combinations of the form

$$p(\sigma) = \sum_\alpha w_\alpha p_\alpha(\sigma), \qquad w_\alpha \geq 0, \quad \forall\alpha, \qquad \sum_\alpha w_\alpha = 1, \tag{5.68}$$

where the $p_\alpha(\sigma)$ are different coexisting equilibrium states. Indeed, it is clear that each such combination has the same value of the free energy per spin of any of the $p_\alpha(\sigma)$. It is known that a set defined in this way can be obtained as the space spanned by the convex combinations of its vertices, that is, of elements that can only be obtained with a combination in which $w_\alpha = \delta_{\alpha\alpha_0}$ for some α_0. In the space of equilibrium states, these vertices are known as **pure states**. Then the generic equilibrium state is obtained by assigning a weight to each pure state. Weights are nonnegative, and sum to 1. This is the generalization of the situation we encountered at the liquid-vapor coexistence, where a fraction of the system was in the liquid phase and the remainder in the vapor one. It can be shown that the pure states, and only they, have the **clustering property**, that is, that faraway spins tend to be independent:

$$\lim_{|\boldsymbol{r}|\to\infty} \langle \sigma_{r+r_0}\sigma_{r_0}\rangle = \langle \sigma_{r+r_0}\rangle \langle \sigma_{r_0}\rangle, \qquad \forall \boldsymbol{r}_0. \tag{5.69}$$

If the coexistence stems from symmetry breaking, the symmetry transformations g map each pure state into a pure state (which can be the same or a different one).

5.9 Peierls's Argument

It is not obvious a priori that statistical mechanics can describe spontaneous symmetry breaking. Indeed, in a finite system, the Boltzmann factor $e^{-\mathcal{H}(\sigma)/k_B T}$ is an analytic function of the temperature and of the other parameters appearing in the Hamiltonian, and thus the partition function $Z = \sum_\sigma e^{-\mathcal{H}(\sigma)/k_B T}$, being the sum of a finite number of analytic functions, is itself analytic.

However, in order to describe the thermodynamic properties of a system, we have to take the thermodynamic limit, in which the size of the system goes to infinity, with fixed

values of the temperature and of the other intensive parameters such as the magnetic field. For example, in the two-dimensional Ising model with $h = 0$, one can show that one obtains in this limit different values of the magnetization per spin m for $h = 0$, provided that the temperature is low enough and the boundary conditions are suitably chosen. This can be proven by means of an argument due to R. Peierls [Peie34]. This was historically important, because it definitely showed that statistical mechanics could describe phase transitions without ad hoc hypotheses. Peierls's argument can easily be generalized to more than two dimensions. Here, I describe it without dwelling excessively on rigor—the argument expounded in this fashion can, however, easily be made rigorous. This section therefore also serves as an introduction to the problems and methods of a rigorous statistical mechanics.

Let us then consider an Ising model defined over a square lattice with $L \times L$ spin. A spin variable $\sigma_i = \pm 1$ is associated with each lattice site i. We denote by σ the microstate of the system, defined by the collection (σ_i) of the values of these variables, one for each site i. The Hamiltonian has the usual Ising form,

$$\mathcal{H}(\sigma) = -\sum_{<i,j>} J\sigma_i\sigma_j - \sum_i h\sigma_i, \tag{5.70}$$

where the sum over $<i,j>$ runs over all pairs of nearest-neighbor sites. When $h = 0$, this Hamiltonian is invariant with respect to spin inversions defined by

$$\sigma \longrightarrow \mathsf{I}\sigma, \qquad \mathsf{I}\sigma_i = -\sigma_i, \quad \forall i. \tag{5.71}$$

Spontaneous symmetry breaking occurs when, in the thermodynamic limit, one can have different states of equilibrium in correspondence with the same temperature value T.

In order to clarify our ideas, let us assume that we impose on the system the boundary conditions $+$, in which the spins external to the system are directed up. In this situation, we expect that there will be a slight prevalence of up spins in the entire finite system at any temperature. On the other hand, if we impose the boundary conditions $-$, in which the spins external to the system are directed down, we will have a prevalence of down spins. The problem is whether this prevalence due to the boundary conditions persists in the thermodynamic limit.

Let us consider the spin placed at the origin, and let us denote by p the probability that it is equal to $+1$, with the boundary conditions $+$. Symmetry breaking occurs if p remains strictly larger than $1/2$ in the thermodynamic limit. The symmetry of the Hamiltonian guarantees in fact that p is also equal to the probability that the spin at the origin is equal to -1 with the boundary conditions $-$. Therefore, if $p > 1/2$, there will be two different equilibrium states at the same temperature T, which are selected by the boundary conditions.

Let us now evaluate p with the boundary conditions $+$ and show that, if the temperature is low enough, one has $p > 1/2$. Let us consider an arbitrary configuration σ of the spins of the system. We can associate a configuration Γ of boundaries between up and down spins with each such configuration. When the boundary conditions are fixed, this correspondence is such that each σ corresponds to one Γ and vice versa.

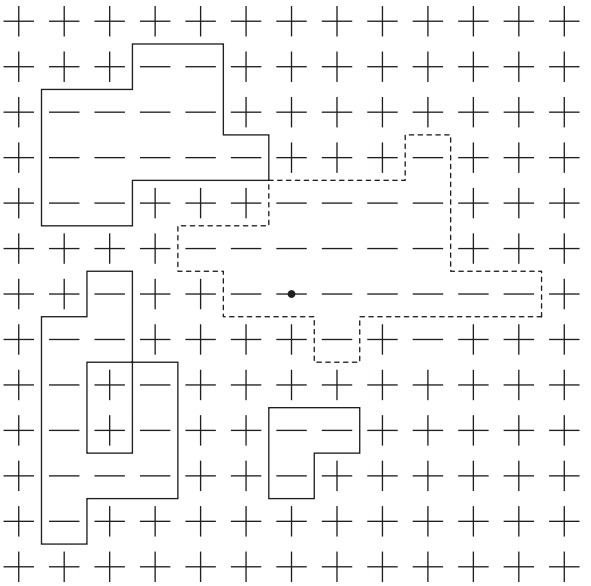


FIGURE 5.6. Configuration of boundaries associated with a spin configuration. The spin at the origin is marked by a black dot.

Let us focus our attention on the spin at the origin (denoted by a black dot in figure 5.6), and divide the configurations into two sets: the set Σ^+ contains all σ configurations in which this spin is equal to $+1$, and the set Σ^- contains all the ones in which the spin at the origin is equal to -1. It is possible to associate a precise configuration $\sigma^+ \in \Sigma^+$ with each configuration $\sigma^- \in \Sigma^-$, by flipping the direction of all the spins inside the smallest boundary that also contains the origin within it. In the case shown in figure 5.6, it is the boundary marked by the dotted line. The same configuration σ^+ can obviously be obtained starting from several different configurations $\sigma^- \in \Sigma^-$.

We then have

$$p = \frac{1}{Z} \sum_{\sigma \in \Sigma^+} \mathrm{e}^{-\mathcal{H}(\sigma)/k_\mathrm{B}T}, \tag{5.72}$$

where Z is the partition function, and the sum runs over all the configurations belonging to Σ^+. We have analogously

$$q = 1 - p = \frac{1}{Z} \sum_{\sigma \in \Sigma^-} \mathrm{e}^{-\mathcal{H}(\sigma)/k_\mathrm{B}T}. \tag{5.73}$$

It is clear that the energy of a configuration σ is equal to the energy E_0 of the ground state ($\sigma_i = +1, \forall i$) plus the contribution of the boundaries. Since each pair of antiparallel nearest-neighbor spins contributes $+2J$ to energy, one has

$$\mathcal{H}(\sigma) = E_0 + 2J\,|\Gamma|\,, \tag{5.74}$$

where $|\Gamma|$ is the total length of the boundaries that appear in σ. Given a configuration $\sigma \in \Sigma^+$, we denote by $\Psi(\sigma)$ the set of all configurations $\sigma' \in \Sigma^-$ that are mapped to σ

by removing the smallest boundary that contains the origin. We denote the removed boundary by $\gamma_\sigma(\sigma')$ and its length with $\left|\gamma_\sigma(\sigma')\right|$. We then have

$$p-q=\frac{1}{Z}\sum_{\sigma\in\Sigma^-}\mathrm{e}^{-\mathcal{H}(\sigma)/k_\mathrm{B}T}\left[1-\sum_{\sigma'\in\Psi(\sigma)}\mathrm{e}^{-2J\left|\gamma_\sigma(\sigma')\right|/k_\mathrm{B}T}\right]. \tag{5.75}$$

We now show that the sum over σ' that appears on the right-hand side is certainly smaller than 1 in the thermodynamic limit, if the temperature is low enough. Let us collect the configurations σ' according to the length $\ell=\left|\gamma_\sigma(\sigma')\right|$ of the smallest boundary that contains the origin. Obviously, $\ell\geq 4$. Let us now evaluate the number $g(\ell)$ of configurations σ' that correspond to a certain value of ℓ. Let us assume that the boundary starts from a point i. The first segment can go in four directions. In subsequent steps, there will be no more than three possible directions, since the boundary cannot retrace its steps. Thus, if we neglect the constraint that the path closes, the total number of paths of length ℓ that leave from and return to i is certainly smaller than $4\times 3^{\ell-1}=(4/3)3^\ell$. The starting point i, however, can be arbitrarily chosen from the ℓ boundary points, and the boundary itself can be traversed in two directions. The number of different boundaries is therefore certainly smaller than $(4/3)(1/2\ell)3^\ell=(2/3\ell)3^\ell$. Since the origin could be at any point inside the boundary, one needs to multiply by the area included in the boundary, which is certainly smaller than or equal to the area of a square whose side is $\ell/4$. We thus obtain

$$g(\ell)<\frac{2}{3\ell}3^\ell\left(\frac{\ell}{4}\right)^2=\frac{\ell}{24}3^\ell. \tag{5.76}$$

One should note that in following this line of reasoning, we omitted the constraints imposed by the presence of other boundaries determined by the configuration $\sigma\in\Sigma^+$ to which σ' is mapped—we are therefore (significantly) overestimating the number of configurations σ'.

To conclude,

$$\mathfrak{S}=\sum_{\sigma'\in\Psi(\sigma)}\mathrm{e}^{-2J\gamma_\sigma(\sigma')/k_\mathrm{B}T}<\sum_{\ell=4}^{\infty}\frac{\ell}{24}w^\ell, \tag{5.77}$$

where

$$w=3\mathrm{e}^{-2J/k_\mathrm{B}T}. \tag{5.78}$$

Therefore

$$\mathfrak{S}<\frac{1}{24}\sum_{n=2}^{\infty}(2nw^{2n})=\frac{w^4(2-w^2)}{12(1-w^2)^2}. \tag{5.79}$$

When $w<w_0\simeq 0.86976$, the expression on the right-hand side is certainly smaller than 1. This occurs for temperatures below

$$T_0=2J/\left[k_\mathrm{B}\ln(3/w_0)\right]\simeq 1.6153\,J/k_\mathrm{B}. \tag{5.80}$$

We therefore see that $p - q > 0$ for temperatures below T_0, and that therefore spontaneous symmetry breaking occurs, and the corresponding critical temperature is higher than T_0. It is obviously possible to obtain a better bound on T_c with little effort. The exact result, for the two-dimensional Ising model, is $T_c = 2.269\ldots\, J/k_B$.

Exercise 5.5 Generalize Peierls's argument to $d = 3$ dimensions.

5.10 The One-Dimensional Ising Model

The preceding argument does not hold in one dimension. It is possible in effect to show that the Ising model in one dimension does not exhibit spontaneous symmetry breaking. The most convincing line of reasoning entails the exact solution of the model, but it is possible to show, by means of an argument due to Landau, that there cannot be discrete symmetry breaking in a one-dimensional system with short-range interactions. This also holds for continuous symmetry. In this section, we describe a (nonrigorous) version of Landau's argument and we then arrive at the exact solution of the one-dimensional Ising model.

5.10.1 The Landau Argument

At the temperature $T = 0$, a one-dimensional Ising system will be in one of the two states of minimal energy—for example, in the one in which all spins are aligned in the $+$ direction. The problem is whether this state can persist in the presence of thermal fluctuations at $T > 0$. Let us assume that we switch the direction of a certain number of consecutive spins—we then obtain an excited state, with an energy that is $2J$ above the minimum. The number of states that have this energy is proportional to the total number of spins N, because we can arrange each boundary between $+$ and $-$ spins in N different fashions. Thus, as soon as $k_B T \ln N > 2J$, reversing a macroscopic number of spins leads to an increase in the entropy proportional to $\ln N$ and therefore to an advantage in free energy. But for $N \to \infty$, this will be true for any positive value of T. This argument can immediately be generalized to all cases of short-range interaction.

Exercise 5.6 Let us suppose that the interaction is long range, so that the coupling constant J_{ij} between the i spin and the j spin decays as $|i-j|^{-\alpha}$, where α is a positive constant. Show that Landau's argument does not rule out a transition if $\alpha < 2$.

5.10.2 Solution of the One-Dimensional Ising Model

The model is defined by the Hamiltonian

$$\mathcal{H}(\sigma) = -J \sum_{i=1}^{N} \sigma_i \sigma_{i+1}, \tag{5.81}$$

where we have imposed periodic boundary conditions:

$$\sigma_{i+N} = \sigma_i, \qquad \forall i. \tag{5.82}$$

The partition function is expressed by

$$Z = \sum_{\sigma} \mathrm{e}^{-\mathcal{H}(\sigma)/k_{\mathrm{B}}T} = \sum_{\{\sigma_i,\ldots,\sigma_N\}} \prod_{i=1}^{N} \mathrm{e}^{K\sigma_i\sigma_{i+1}}, \tag{5.83}$$

where $K = J/k_{\mathrm{B}}T$. By introducing the **transfer matrix** $\mathsf{T} = (T_{\sigma\sigma'})$, defined by

$$T_{\sigma\sigma'} = \mathrm{e}^{K\sigma\sigma'}, \tag{5.84}$$

we obtain the following expression of Z:

$$\begin{aligned} Z &= \sum_{\sigma_1}\sum_{\sigma_2}\cdots\sum_{\sigma_N} T_{\sigma_1\sigma_2}T_{\sigma_2\sigma_3}\cdots T_{\sigma_{N-1}\sigma_N}T_{\sigma_N\sigma_1} \\ &= \mathrm{Tr}\,\mathsf{T}^N. \end{aligned} \tag{5.85}$$

The transfer matrix T has two eigenvalues, $t_+ = 2\cosh K$ and $t_- = 2\sinh K$, with $t_+ > t_-$. Therefore for $N \gg 1$,

$$\mathrm{Tr}\,\mathsf{T}^N = t_+^N + t_-^N \simeq t_+^N. \tag{5.86}$$

In the thermodynamic limit, therefore, the free energy per spin is given by

$$-\frac{f}{k_{\mathrm{B}}T} = \lim_{N\to\infty}\frac{\ln Z}{N} = \ln t_+ = \ln 2\cosh K. \tag{5.87}$$

This expression does not exhibit singularities for $0 < K < \infty$, and the model therefore does not exhibit phase transitions at finite temperature.

Exercise 5.7 Consider the spin correlation function

$$C_{ij} = \langle\sigma_i\sigma_j\rangle - \langle\sigma_i\rangle\langle\sigma_j\rangle$$

in the one-dimensional Ising model with a vanishing applied field. Show that one has

$$\lim_{N\to\infty} C_{ij} = (\tanh K)^{|i-j|}.$$

This expression decays as $\mathrm{e}^{-|r_i - r_j|/\xi}$, where ξ is called the **coherence length** and is given by

$$\xi = a_0/\left|\ln\tanh K\right|,$$

where a_0 is the lattice constant.

Exercise 5.8 Show that, in the presence of a nonvanishing magnetic field h, in the one-dimensional Ising model, one has

$$-\frac{f}{k_{\mathrm{B}}T} = \lim_{N\to\infty}\frac{\ln Z}{N} = \ln\left[\mathrm{e}^K\cosh\lambda + \sqrt{\mathrm{e}^{2K}\sinh^2\lambda + \mathrm{e}^{-2K}}\right],$$

where $\lambda = h/k_{\mathrm{B}}T$.

Exercise 5.9 Obtain the same result as in the previous exercise by the following method, introduced by S. Lifson [Lifs64]. Let us consider the system with $\sigma_0 = +1$, and

with σ_N free. Then, in analogy with the Peierls argument, the spin configurations σ are univocally defined by the configurations Γ of the boundaries—in other words, of nearest-neighbor pairs with opposite spin values. The system is split into n intervals, of length $\ell_i \geq 1$, such that sites belonging to the same interval have the same value of the spin.

1. Write the canonical partition function as a function of the σ_is.
2. Move to the corresponding grand canonical partition function by introducing the spin fugacity z.
3. Evaluate the grand partition function as a function of z.
4. Evaluate $\langle N \rangle$ and prove that, in order to go to the thermodynamic limit $N \to \infty$, it is necessary that $z \to z^*$, where z^* is the value of z closest to the origin for which Z admits a singularity. Evaluate z^*.
5. Show that $-f/k_B T = \lim_{N\to\infty} \ln Z/N = -\ln z^*$, and compare the result with what was already obtained.

Exercise 5.10 (The Chinese philosophers' problem) Let us consider N Chinese philosophers sitting at a circular table. Each philosopher can either meditate (state "-1") or eat (state "$+1$"). In order to eat he needs two chopsticks: chopsticks are shared between the two neighboring philosophers, one on the left and one on the right. Thus if a philosopher is eating, his left and right neighbors cannot be eating at the same time. We wish to evaluate the number of possible configurations of the system of N philosophers.

1. Show that the question reduces to the following problem: one considers N sites on a circle, each carrying a spin variable $\sigma_i \in \{-1, +1\}$. The total number of configurations is 2^N. We wish to evaluate the number $\mathcal{N}_N$ of states in which one cannot find two spins $+1$ one after the other.
2. Evaluate explicitly $\mathcal{N}_N$ for $N = 2, 3, 4, 5$.
3. Check that the calculated $\mathcal{N}_N$ verify the relation $\mathcal{N}_N = \mathcal{N}_{N-1} + \mathcal{N}_{N-2}$, which defines the Fibonacci sequence. The "standard" Fibonacci sequence satisfies this relation and the initial conditions $\mathcal{N}_0 = \mathcal{N}_1 = 1$. In our case the initial conditions will be different.
4. Let us denote by $\mathcal{N}_N(\sigma)$ the number of acceptable configurations in which $\sigma_N = \sigma$, $\sigma \in \{-1, +1\}$. Show that one then has

   $\mathcal{N}_N = \mathcal{N}_N(+1) + \mathcal{N}_N(-1)$.

   Evaluate $\mathcal{N}_N(+1)$ as a function of $\mathcal{N}_{N-1}$ and $\mathcal{N}_{N-1}(+1)$, and analogously for $\mathcal{N}_N(-1)$, and show that one can derive the Fibonacci recursion relation.
5. Show that the problem reduces to counting the minimal-energy states of a one-dimensional Ising model defined by the Hamiltonian

$$\mathcal{H}(\{\sigma_i\}) = \sum_{i=1}^{N} \left[-J\sigma_i\sigma_{i+1} - \frac{h}{2}(\sigma_i + \sigma_{i+1}) \right],$$

with periodic boundary conditions ($\sigma_{i+N} = \sigma_i$) and with

$J = -1; \qquad h = -2.$

Evaluate the corresponding ground-state energy $E_0(N)$.

6. Assuming we are able to evaluate the canonical partition function

$$Z_N(\beta) = \sum_{\{\sigma_i\}} \mathrm{e}^{-\beta H(\{\sigma_i\})},$$

show that one has

$$\mathcal{N}_N = \lim_{\beta \to \infty} \mathrm{e}^{\beta E_0(N)} Z_N(\beta).$$

7. Given the expression for the canonical partition function of the one-dimensional Ising model,

$$Z_N(\beta) = t_+^N + t_-^N,$$

where $t_\pm$ are the two eigenvalues of the transfer matrix of the system, expressed as functions of β, J, and h, check the values of $\mathcal{N}_N$ previously obtained.

5.11 Duality

The critical temperature of the two-dimensional Ising model was found in 1941 by H. A. Kramers and G. H. Wannier [Kram41] before the exact solution to the model, which is due to L. Onsager [Onsa44]. Kramers and Wannier's argument introduces a transformation—called **duality**—which maps a high-temperature Ising model into an analogous low-temperature model. This transformation is of great importance and can be generalized to other statistical models, even though, in the large majority of cases, the transformed model is not identical to the initial one. We deal with it only in the case of the Ising model in $d = 2$, by means of an argument that is simpler than the original one. This argument also allows us to introduce the high- and low-temperature expansions, which are used to obtain information about the behavior of models that cannot be solved exactly.

Let us therefore consider the expression of the partition function of a two-dimensional Ising model at temperature T in a vanishing external field:

$$Z(K) = \sum_{\sigma} \exp\left(K \sum_{<ij>} \sigma_i \sigma_j \right), \tag{5.88}$$

where $K = J/k_\mathrm{B} T$ and the sum over σ runs over all the 2^N spin configurations and that over $<ij>$ runs over the $2N$ pairs of nearest-neighbor sites. Since

$$\mathrm{e}^{K\sigma\sigma'} = \cosh K + \sigma\sigma' \sinh K = \cosh K \left(1 + \sigma\sigma' \tanh K\right), \tag{5.89}$$

Z can also be written

$$Z(K) = (\cosh K)^{2N} \sum_{\sigma} \prod_{<ij>} \left(1 + \sigma_i \sigma_j \tanh K\right). \tag{5.90}$$

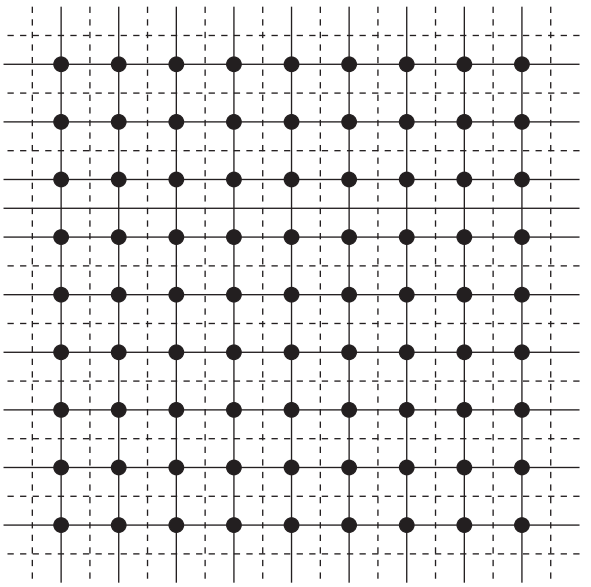


FIGURE 5.7. Dual lattices. The spins denoted by black dots are arranged in a square lattice, marked by the full lines. These lines define plaquettes in whose centers are placed the spins of the dual lattice. The dual lattice is marked by the dashed lines.

We can now expand the expression on the right-hand side into a power series of $t = \tanh K$. Since $K \propto 1/T$, this yields a high-temperature expansion. The terms of this expansion can be represented by diagrams. One marks a number of nearest-neighbor bonds on the lattice and associates a factor $t\sigma_i\sigma_j$ to each bond of the diagram that connects site i with site j. Each bond can appear at most once. We thus obtain

$$Z(K) = (\cosh K)^{2N} \sum_{\sigma} \sum_{\mathcal{G}} t^{|\mathcal{G}|} \prod_i \sigma_i^{n_i(\mathcal{G})}, \tag{5.91}$$

where the sum runs over all the diagrams $\mathcal{G}$ that can be drawn on the lattice, $|\mathcal{G}|$ is the number of bonds that appear in diagram $\mathcal{G}$, and for each site i, $n_i(\mathcal{G})$ is the number of bonds that contain the i site in diagram $\mathcal{G}$. When we sum all the spin configurations, all the diagrams $\mathcal{G}$ in which $n_i(\mathcal{G})$ is odd for one or more i sites will result in a zero contribution, whereas each of the other diagrams yields a contribution equal to $2^N(\cosh K)^{2N} t^{|\mathcal{G}|}$. Therefore,

$$Z(K) = 2^N(\cosh K)^{2N} \sum_{\mathcal{G}}{}' t^{|\mathcal{G}|}, \tag{5.92}$$

where the sum runs over all the diagrams in which the number of bonds that contain each i site is even ($n_i(\mathcal{G}) = 0, 2, 4, \forall i$).

We can now interpret the diagrams $\mathcal{G}$ that contribute to Z as configurations Γ of boundaries between opposite spins in an Ising model defined on a **dual lattice**. To define the dual lattice, consider fig. 5.7. The bounds connecting nearest-neighbor spins in the model divide the plane into **plaquettes**, that is, squares whose corners are occupied by the spins

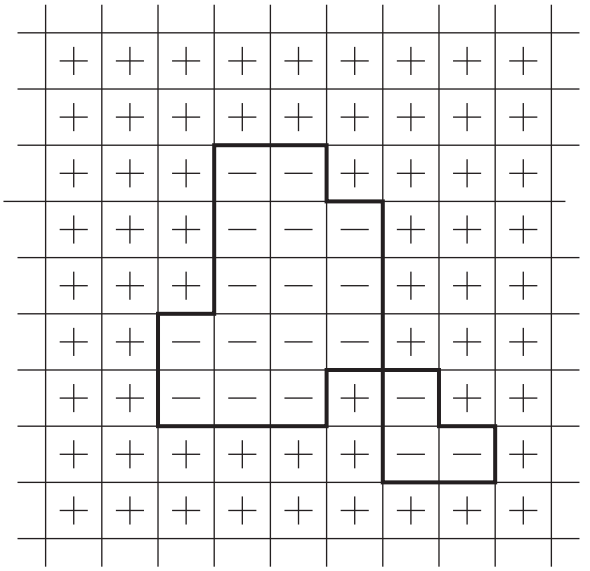


FIGURE 5.8. Diagram $\mathcal{G}$ in the high-temperature expansion of the Ising model interpreted as a configuration Γ of boundaries in the Ising model defined over the dual lattice.

σ_i, and whose sides are the bonds that connect them. Associate a spin $\hat{\sigma}_I$ to each plaquette I. Then each diagram $\mathcal{G}$ that contributes to Z is made of loops that delimit a certain numbers of plaquettes in their interiors. These loops separate regions where the spins $\hat{\sigma}_I$ in the interior are opposite to those in the exterior.

In figure 5.8, the original model is defined on the vertices of the lattice. A diagram $\mathcal{G}$ that contributes to Z is drawn with a thick line. This diagram can be interpreted as a configuration Γ of boundaries in a spin model defined over the dual lattice—where an Ising spin variable is associated with each plaquette.

We thus obtain

$$\sum_{\mathcal{G}}{}' t^{|\mathcal{G}|} = \sum_{\Gamma} t^{|\Gamma|}, \tag{5.93}$$

where $|\Gamma|$ is the total length of the boundaries that appear in Γ.

We now consider the expression of the partition function of the Ising model over the dual lattice, and we denote its temperature with $\hat{T}$. We impose the $+$ boundary conditions. One then has

$$Z(\hat{K}) = \mathrm{e}^{2N\hat{K}} \sum_{\Gamma} \mathrm{e}^{-2\hat{K}|\Gamma|}, \tag{5.94}$$

where $\hat{K} = J/k_{\mathrm{B}}\hat{T}$. By comparing this with the expression of $Z(K)$, we obtain the relation

$$t = \tanh K = \mathrm{e}^{-2\hat{K}}. \tag{5.95}$$

This is the duality relation for temperatures. It is a duality relation in the sense that if we define D as the mapping $K \longrightarrow \hat{K}$, we have $\mathsf{D}^2 K = \mathsf{D}\hat{K} = K$, that is, the D transformation is

equal to its inverse. Let us note that for $K \to 0$, one has $\hat{K} \to \infty$, and that duality therefore maps high-temperature systems to low-temperature systems, and vice versa. By making use of the relation (5.95), we obtain

$$Z(K) = \mathrm{e}^{2N\hat{K}} \frac{1}{(2\cosh^2 K)^N} Z(\hat{K}) = (\sinh 2\hat{K})^N Z(\hat{K}). \tag{5.96}$$

If the Ising model admits a single transition at $h = 0$, it must occur at a point at which $\hat{K} = K$. We thus obtain an equation for the critical value K_c,

$$\tanh K_c = \mathrm{e}^{-2K_c}, \tag{5.97}$$

which yields

$$K_c = \frac{1}{2} \ln\left(1 + \sqrt{2}\right) = 0.44069\ldots, \tag{5.98}$$

which corresponds to $k_B T_c = 2.2692\ldots J$.

Exercise 5.11 (Duality in the three-dimensional Ising model) By considering the expansion of the partition function of the three-dimensional Ising model as a function of the distribution Γ of boundaries, show that the dual of this model is defined by the Hamiltonian

$$\mathcal{H}(\hat{\sigma}) = -J \sum_{\mathcal{P}} \prod_{I \in \mathcal{P}} \hat{\sigma}_I, \tag{5.99}$$

where the sum runs over all plaquettes $\mathcal{P}$ in the dual lattice, and the product runs over all spins that belong to each plaquette. Find the relation between the temperatures of the Ising model and the corresponding dual model.

Note that, since the dual of the Ising model is not identical to the initial model, this argument does not provide us with information about the critical temperature for $d = 3$.

5.12 Mean-Field Theory

In order to qualitatively calculate the phase diagram for systems in statistical mechanics, one often resorts to the **mean-field theory**. This approximated theory is extremely useful, and is the first instrument one resorts to when exploring new models. I present it in the simplest case—in other words, in the Ising model.

Let us suppose that we mentally isolate a spin (which we denote by 0) from its environment, in the Ising model. If its state changes, passing for instance from $\sigma_0 = +1$ to $\sigma_0' = -1$, while all the rest of the environment remains constant, the energy of the system varies by

$$\Delta\mathcal{H} = 2h + 2 \sum_{j \in \text{n.n.}(0)} J\sigma_j = -\left(h + \sum_{j \in \text{n.n.}(0)} J\sigma_j\right) \Delta\sigma_0, \tag{5.100}$$

where $\Delta\sigma_0 = \sigma_0' - \sigma_0$ and the sum runs over the nearest-neighbor sites of 0. This is the same variation that one would obtain for a paramagnet immersed in a field whose value is equal to

$$h_{\text{eff}}(\sigma) = h + \sum_{j\in\text{n.n.}(0)} J\sigma_j. \tag{5.101}$$

However, the spins close to the one being considered fluctuate. As a first approximation, we can assume that their action on the spin 0 is approximately given by the mean value of this field—hence, the name mean-field theory. We thus obtain

$$h_{\text{eff}} = \langle h_{\text{eff}}(\sigma)\rangle = h + \sum_{j\in\text{n.n.}(0)} J\langle\sigma_j\rangle. \tag{5.102}$$

Let us calculate the mean value of σ for an isolated spin, subjected to an arbitrary external field h. One has

$$\langle\sigma\rangle = \frac{1}{z_0}\sum_\sigma \sigma\, \mathrm{e}^{h\sigma/k_\mathrm{B}T}, \qquad z_0 = \sum_\sigma \mathrm{e}^{h\sigma/k_\mathrm{B}T}. \tag{5.103}$$

A simple calculation yields

$$\langle\sigma\rangle = \tanh\left(\frac{h}{k_\mathrm{B}T}\right). \tag{5.104}$$

By exploiting this result, and substituting h with h_{eff} we obtain

$$\langle\sigma_0\rangle = \tanh\left(\frac{h_{\text{eff}}}{k_\mathrm{B}T}\right) = \tanh\left[\frac{1}{k_\mathrm{B}T}\left(h + \sum_{j\in\text{n.n.}(0)} J\langle\sigma_j\rangle\right)\right]. \tag{5.105}$$

We can now observe that, in the ordered phase, we expect that the mean value of σ_i will be the same for all spins. This observation is not as trivial as it seems—if in fact we were to change the sign of the interaction J, all the observations we made so far would be just as valid, except for this one, because we would instead expect that the $\langle\sigma_i\rangle$ assume opposite signs on each of the intertwined sublattices in the ordered phase. More generally, the most difficult point of the mean-field theory is that one needs to make a good guess about the type of order one expects. In our case, we posit

$$\langle\sigma_i\rangle = m, \quad \forall i, \tag{5.106}$$

from which we obtain an equation for m:

$$m = \tanh\left(\frac{h + \zeta J m}{k_\mathrm{B}T}\right), \tag{5.107}$$

where ζ is the **coordination number** of the lattice, that is, the number of nearest neighbors of a given site.

It is easier to discuss this equation in the form

$$\tanh^{-1} m = \frac{h + \zeta J m}{k_\mathrm{B}T}. \tag{5.108}$$

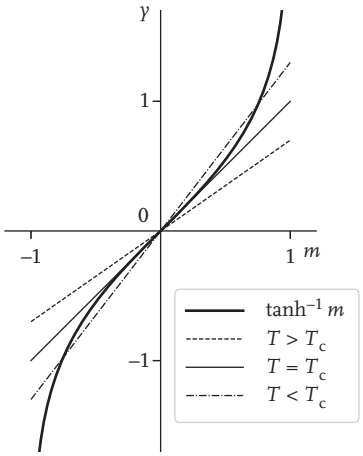


FIGURE 5.9. Graphical solution of the mean-field eq. (5.108) for $h = 0$. The magnetization per spin, m, is obtained at the crossing of the curve $y = \tanh^{-1} m$ with the line $y = \zeta J m / k_B T$.

Then the possible solutions for m are the intersections of the curve $y = \tanh^{-1} x$ with the lines $y = (h + \zeta J x)/k_B T$. The graphical solution for $h = 0$ is shown in fig. 5.9. The derivative of $\tanh^{-1} x$ assumes its minimum value at the origin, where it is equal to 1. Thus a line passing through the origin with slope less than or equal to 1 has only one intersection (at the origin) with the curve. This corresponds to $k_B T > \zeta J$. On the other hand, if the slope is larger than 1, there are three intersections: one at the origin, and two symmetrical ones at $\pm m_0(T)$. If the magnetic field h does not vanish, but is small, the situation will be similar, but the intersections will be slightly displaced.

We can then identify the critical temperature T_c with $\zeta J/k_B$. Which of the three intersections for $T < T_c$ is the "physical" solution? Thermodynamics tells us that we must choose the one in which the free energy $F = E - TS - hM$ is minimal. In our case, we are consistent if we estimate the entropy S from the corresponding expression valid for independent spins:

$$S = -Nk_B \left[\left(\frac{1-m}{2} \right) \ln \left(\frac{1-m}{2} \right) + \left(\frac{1+m}{2} \right) \ln \left(\frac{1+m}{2} \right) \right]. \tag{5.109}$$

We then have, for small values of m,

$$\frac{F}{N} = -\frac{\zeta J}{2} m^2 - k_B T \left(\ln 2 - \frac{m^2}{2} - \cdots \right) - hm. \tag{5.110}$$

By making use of this expression, it is easy to see that the "physical" solution is the intersection away from the origin, of the same sign as the external field h. Letting $h \to 0$, we obtain the spontaneous magnetization m_0 as a function of T, represented by the curve shown in figure 5.10.

We can also calculate the specific heat a vanishing external field:

$$\begin{aligned} C_h(T) &= T \left(\frac{\partial S}{\partial T} \right)_h = T \frac{\mathrm{d}S}{\mathrm{d}m} \left(\frac{\partial m}{\partial T} \right)_h \\ &= \begin{cases} Nk_B m^2 [(T/T_c)^2/(1-m^2) - T/T_c]^{-1}, & \text{for } T < T_c, \\ 0, & \text{for } T \geq T_c. \end{cases} \end{aligned} \tag{5.111}$$

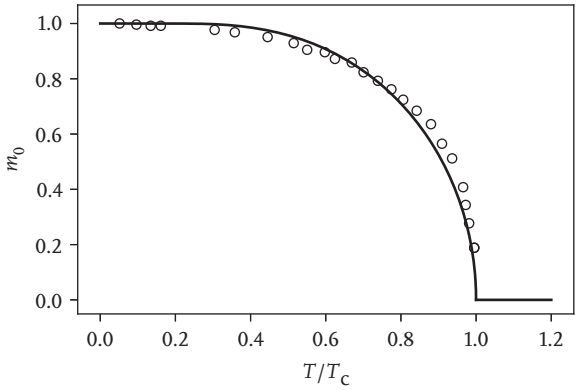


FIGURE 5.10. Spontaneous magnetization per spin, $m_0(T)$, in the Ising model, as a function of $T/T_c = k_B T/\zeta J$, in the mean-field approximation. Data on nickel by P. Weiss and R. Forrer [Weis26].

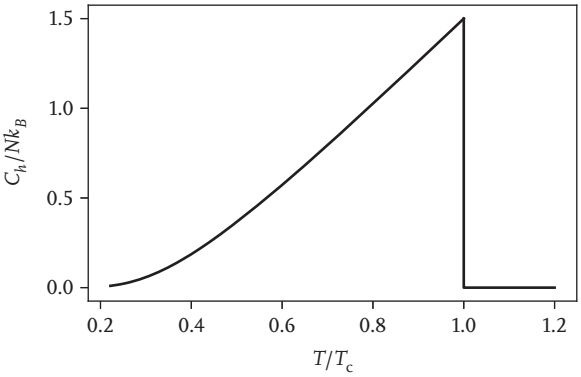


FIGURE 5.11. Specific heat C_h per spin (in units of k_B) for $h = 0$ in an Ising model, as a function of $T/T_c = k_B T/\zeta J$, in the mean-field approximation.

This behavior is shown in fig. 5.11. We have $\lim_{T\to T_c^-} C(T) = 3Nk_B/2$. Thus in mean-field theory the specific heat exhibits a discontinuity at T_c, since it identically vanishes above the critical temperature, at least in our approximation. The behavior of actual systems is qualitatively similar, but it presents a more complex singularity than a simple discontinuity (see, for example, fig. 5.12). The phase diagram is particularly simple in the (h, T) plane. The coexistence curve is simply the line $(0, 0)$–$(0, T_c)$. The magnetization m is subject to a discontinuity along this line, going from $m_0(T)$ for positive h to $-m_0(T)$ for negative h.

If we take at face value the result of mean-field theory for the liquid-gas transition, we see that phase coexistence is characterized by the discontinuity in the entropy density, that is, in the first derivative of the free energy density with respect to the temperature, while the critical point exhibits a discontinuity in the specific heat, that is, in its second derivative. In 1933, P. Ehrenfest [Ehre33] generalized this observation to define an n-order phase transition if there is a discontinuity of the nth derivative of the free energy density at the transition point, while all derivatives of lower order are continuous. Thus the liquid-gas transition would be a first-order one, and the critical point a second-order one.

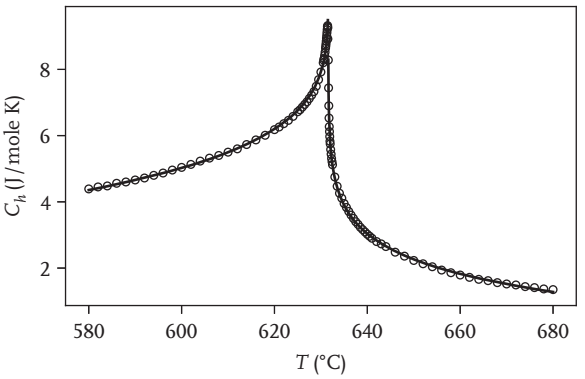


FIGURE 5.12. Magnetic contribution to the specific heat of nickel close to the Curie temperature. The curve is a fit to the expression $C_h = (A/\alpha)\,|T - T_c|^{-\alpha} + K$, separately below and above $T_c = 631.58$°C. Data by Connelly et al. [Conn71].

However, many examples of phase transitions do not fit into this scheme, and therefore this classification is now considered obsolete.

5.13 Variational Principle

The reasoning we reproduced in the preceding section somehow lacks consistency. Once we had introduced the concept of mean field, we saw that there can be several thermodynamic states (defined by different values of magnetization per spin m) that are solutions of the corresponding equations. To resolve this ambiguity, we recalled the variational principle of free energy, approximating entropy with its expression as obtained for the paramagnet.

In this section, we prove that mean-field theory can be obtained from a variational principle, which allows us to derive both the self-consistency equations we described and the form of the free energy we exploited. This also allows us to clarify the hypotheses that provide the foundation for mean-field theory.

The starting point is **Jensen's inequality**, discussed in sec. B.2. When applied to the exponential, it yields

$$\left\langle \exp f(x) \right\rangle \geq \exp \left\langle f(x) \right\rangle , \tag{5.112}$$

where $f(x)$ is an arbitrary real-valued function of the random variable x.

This inequality suggests that we use a variational principle when calculating the free energy. Let us introduce a "trial Hamiltonian" $\mathcal{H}_0(\sigma)$ (arbitrary for the time being). We can then write

$$\begin{aligned} Z &= \sum_{\sigma} \exp\left[-\frac{\mathcal{H}(\sigma)}{k_B T}\right] = Z_0 \left\langle \exp\left[-\frac{\mathcal{H}(\sigma) - \mathcal{H}_0(\sigma)}{k_B T}\right] \right\rangle_0 \\ &\geq Z_0 \exp\left[-\frac{1}{k_B T} \left\langle \mathcal{H}(\sigma) - \mathcal{H}_0(\sigma) \right\rangle_0\right], \end{aligned} \tag{5.113}$$

where we have defined

$$Z_0 = \sum_\sigma \exp\left[-\frac{\mathcal{H}_0(\sigma)}{k_B T}\right] = \exp\left[-\frac{\langle \mathcal{H}_0\rangle_0 - TS_0}{k_B T}\right], \tag{5.114}$$

$$\langle A(\sigma)\rangle_0 = \frac{1}{Z_0}\sum_\sigma A(\sigma)\ \exp\left[-\frac{\mathcal{H}_0(\sigma)}{k_B T}\right], \tag{5.115}$$

and where S_0, defined in eq. (5.114), is the entropy of the equilibrium state defined by $\mathcal{H}_0$ at the temperature T.

The idea is therefore to look for the maximum of the expression

$$\mathcal{Z} = Z_0 \exp\left[-\frac{1}{k_B T}\langle \mathcal{H} - \mathcal{H}_0\rangle_0\right] \tag{5.116}$$

with respect to the possible choices of the trial Hamiltonian $\mathcal{H}_0$. The easiest choice is a Hamiltonian with independent spins:

$$\mathcal{H}_0(\sigma) = -\sum_i \lambda\, \sigma_i. \tag{5.117}$$

Let us now shift to the logarithms, introducing the true free energy $F = -k_B T \ln Z$ and the trial free energy $\mathcal{F} = -k_B T \ln \mathcal{Z}$. The variational principle then takes the form

$$F \le \mathcal{F}(T, h; \lambda) = \langle \mathcal{H}\rangle_0 - TS_0, \tag{5.118}$$

where the internal energy $\langle \mathcal{H}\rangle_0$ and the entropy S_0 are evaluated by means of the trial Hamiltonian. Since, in this simple case, there is a one-to-one correspondence between the parameter λ and the magnetization m, we can choose m rather than λ as the variational parameter. We thus obtain

$$\langle \mathcal{H}\rangle_0 = -\frac{N\zeta J}{2}m^2 - Nhm; \tag{5.119}$$

$$S_0 = -Nk_B\left[\left(\frac{1-m}{2}\right)\ln\left(\frac{1-m}{2}\right) + \left(\frac{1+m}{2}\right)\ln\left(\frac{1+m}{2}\right)\right]. \tag{5.120}$$

Since $\mathrm{d}S_0/\mathrm{d}m = -Nk_B \tanh^{-1} m$, the variational equation reads

$$-\zeta J m - h + k_B T \tanh^{-1} m = 0, \tag{5.121}$$

which is the one we had obtained in the preceding section. This time, however, we additionally obtain the criterion by which to choose the solutions and, even more importantly, the way to generalize the method—it will be sufficient to choose trial Hamiltonians $\mathcal{H}_0$ made by sums of terms with a few variables that are independent of the others (so as to allow the explicit calculation of the Z_0 and of the average $\langle \mathcal{H}_0\rangle_0$), but that contain a part of the spin interactions. These generalizations of the mean-field theory are called **cluster approximations**.

Exercise 5.12 (Mean-field theory of the Ising antiferromagnet) The Ising antiferromagnet in d dimensions is described by the Hamiltonian

$$\mathcal{H}(\sigma) = \sum_{<ij>} J\sigma_i\sigma_j - \sum_i h\sigma_i, \tag{5.122}$$

where $\sigma = (\sigma_1, \ldots, \sigma_N)$, $\sigma_i = \pm 1$, $\forall i$, and the first sum runs over all pairs $<ij>$ of nearest-neighbor lattice sites. Here J is a positive constant, and the lattice is a simple cubic lattice in d dimensions. As a consequence, the lattice breaks down into two intertwined sublattices, which we denote by $+$ and $-$, such that, if i is a site that belongs to one lattice, its $2d$ nearest neighbors all belong to the other lattice.

Since the Hamiltonian favors antiparallel nearest-neighbor orientations, there is the possibility of spontaneous breaking of the symmetry between the two sublattices.

1. Define a trial free energy for the model, as a function of the magnetization per spin $m_\pm$ of each sublattice: $m_+ = \langle\sigma_i\rangle$, with i belonging to the $+$ sublattice, and analogously for m_-.
2. Write down the self-consistency equations that determine the values of $m_\pm$ as a function of $K = J/k_B T$ and $\lambda = h/k_B T$.
3. Express the critical value $K_c(h)$ of K as a function of the applied magnetic field.
4. Show that the transition is *reentrant* in the (T, h) plane: for a certain interval of h, there are two critical values T_c of T, such that the system exhibits spontaneous symmetry breaking for values of T between them.

5.14 Correlation Functions

In the Ising model, the spin variables σ_i relative to different sites are not independent. A measure of their dependence is given by the correlation function $\mathsf{C} = (C_{ij})$ (cf. sec. A.7), where

$$C_{ij} = \langle\sigma_i\sigma_j\rangle_c = \langle\sigma_i\sigma_j\rangle - \langle\sigma_i\rangle\langle\sigma_j\rangle. \tag{5.123}$$

This quantity vanishes if σ_i and σ_j are independent.

In this section, we will calculate C_{ij} in the context of mean-field theory, using a method due to Ornstein and Zernike [Orns14]. In order to do so, let us rewrite the partition function, expressing it as a function of an interaction matrix $\mathsf{K} = (K_{ij})$, which is a priori arbitrary (but symmetrical), and of an external field $\lambda = (\lambda_i)$, which can vary from point to point:

$$Z = \exp\left(\frac{1}{2}\sum_{ij} K_{ij}\sigma_i\sigma_j + \sum_i \lambda_i\sigma_i\right). \tag{5.124}$$

The physical case corresponds to

$$K_{ij} = \begin{cases} J/k_B T, & \text{if } i \text{ and } j \text{ are nearest neighbors,} \\ 0, & \text{otherwise.} \end{cases} \tag{5.125}$$

$$\lambda_i = h/k_B T, \quad \forall i.$$

We observe that

$$\begin{aligned}\left.\frac{\partial \ln Z}{\partial \lambda_i}\right|_{\text{phys}} &= \langle \sigma_i \rangle = m, \\ \left.\frac{\partial^2 \ln Z}{\partial \lambda_i \, \partial \lambda_j}\right|_{\text{phys}} &= \left\langle \sigma_i \sigma_j \right\rangle_{\text{c}} .\end{aligned} \tag{5.126}$$

The "phys" notation indicates that the derivatives must be calculated with the physical values of K and λ given in eq. (5.125). Let us now evaluate these quantities, using the mean-field expression of the partition function that we derived earlier. We have

$$\ln Z = -\frac{S_0}{k_{\text{B}}} + \frac{1}{2}\sum_{ij} K_{ij} m_i m_j + \sum_i \lambda_i m_i = \ln Z_0 + \sum_i \lambda_i m_i, \tag{5.127}$$

where

$$S_0 = -k_{\text{B}} \sum_i \left[\left(\frac{1-m_i}{2} \right) \ln \left(\frac{1-m_i}{2} \right) + \left(\frac{1+m_i}{2} \right) \ln \left(\frac{1+m_i}{2} \right) \right]. \tag{5.128}$$

Taking the derivative of this expression with respect to λ_i, we obtain

$$\frac{\partial \ln Z}{\partial \lambda_i} = \sum_j \frac{\partial \ln Z_0}{\partial m_j} \frac{\partial m_j}{\partial \lambda_i} + m_i, \tag{5.129}$$

where the local magnetization m_i is given by the mean-field equation:

$$0 = \frac{\partial \ln Z}{\partial m_i} = -\tanh^{-1} m_i + \sum_{j \in \text{n.n.}(i)} K_{ij} m_j + \lambda_i. \tag{5.130}$$

By taking the derivative of this equation with respect to λ_j, we obtain an equation for the correlation function $C_{ij} = \partial m_i / \partial \lambda_j$:

$$\frac{1}{1-m_i^2} C_{ij} = \sum_\ell K_{i\ell} C_{\ell j} + \delta_{ij}. \tag{5.131}$$

We can now substitute the "physical" values of K and λ. Then K_{ij} is translation-invariant, that is, it depends only on the vector $\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$ and $m_i = m$, $\forall i$. Therefore also C_{ij} is translation-invariant, and we can solve the above equation in Fourier space.

We define the Fourier transform according to

$$C(\boldsymbol{k}) = \sum_j \exp\left(-\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}_{ij}\right) C_{ij}, \tag{5.132}$$

and analogously for K. By exploiting the fact that the Fourier transform of a convolution is equal to the product of the Fourier transforms of the convoluted functions (appendix D, eq. (D.20)), eq. (5.131) becomes

$$\frac{1}{1-m^2} \sum_j \exp\left[-\mathrm{i}\boldsymbol{k} \cdot (\boldsymbol{r}_j - \boldsymbol{r}_i)\right] C_{ij} = K(\boldsymbol{k}) C(\boldsymbol{k}) + 1. \tag{5.133}$$

This yields

$$C(\boldsymbol{k}) = \frac{1-m^2}{1-(1-m^2)K(\boldsymbol{k})}. \tag{5.134}$$

We can make this expression more explicit by recalling that, for a simple cubic lattice in d dimensions, one has

$$K(\boldsymbol{k}) = \sum_{j\in \text{n.n.}(i)} \frac{J}{k_\text{B}T} \exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_{ij}\right) = \sum_{\alpha=1}^{d} \frac{J}{k_\text{B}T} 2\cos(k_\alpha a_0), \tag{5.135}$$

where a_0 is the lattice constant. For small values of $k = |\boldsymbol{k}|$, one has

$$K(\boldsymbol{k}) \simeq \frac{J}{k_\text{B}T}\left(\zeta - k^2 a_0^2\right), \tag{5.136}$$

where $\zeta = 2d$ is the coordination number of the lattice. We thus obtain

$$\begin{aligned} C(\boldsymbol{k}) &= (1-m^2)\left[1-(1-m^2)(J/k_\text{B}T)\sum_\alpha 2\cos(k_\alpha a_0)\right]^{-1} \\ &\simeq (1-m^2)\left[1-(1-m^2)(T_\text{c}/T)(1-k^2a_0^2)\right]^{-1}, \end{aligned} \tag{5.137}$$

in which we have substituted $T_\text{c} = \zeta J/k_\text{B}$.

Let us remark that the Fourier transform of the correlation function, evaluated at $\boldsymbol{k} = 0$, is proportional to the magnetic susceptibility. One has in fact

$$C(\boldsymbol{k}{=}0) = \sum_j \frac{\partial\langle\sigma_j\rangle}{\partial\lambda_j} = k_\text{B}T\frac{\partial\langle\sigma_i\rangle}{\partial h} = k_\text{B}T\chi. \tag{5.138}$$

Evaluating this expression for $h = 0$ and $T > T_\text{c}$ (and therefore for $m = 0$), we obtain

$$\chi = \frac{1}{k_\text{B}(T-T_\text{c})}, \tag{5.139}$$

which diverges for $T \to T_\text{c}$. This behavior is called the **Curie-Weiss law** [Weis07], and is shown in fig. 5.13 with data on dysprosium, with an estimated $T_\text{c} = 151$ K.

It is possible to characterize the range of the correlations by means of the **correlation length** ξ, defined by

$$\xi^2 = \frac{1}{2d}\frac{1}{C(\boldsymbol{k}{=}0)}\sum_j r_{ij}^2 C_{ij}, \tag{5.140}$$

where the $1/2d$ factor has been introduced for convenience. It is easy to see that

$$\xi^2 = -\frac{1}{2d}\sum_\alpha \left.\frac{\partial^2 \ln C(\boldsymbol{k})}{\partial k_\alpha^2}\right|_{k=0} = \frac{a_0^2}{1-(T_\text{c}/T)} \propto (T-T_\text{c})^{-1}. \tag{5.141}$$

We thus obtain the important result that the correlation range diverges when $T \to T_\text{c}$. This result is not unexpected, given that the divergence of the susceptibility per spin implies

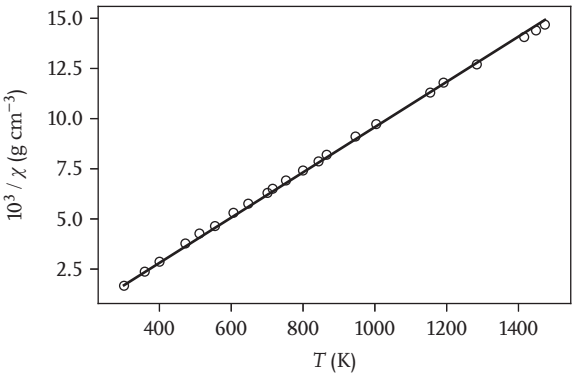


FIGURE 5.13. Inverse susceptibility $1/\chi$ (in cgs units) for Dy as a function of temperature. Notice that the data do not cover the region close to T_c. Adapted from S. Arajs and R. V. Colvin [Araj61], with permission of AIP publishing.

that the number of terms different from zero in the expression of $C(\boldsymbol{k}{=}0) = \sum_j C_{ij}$ must diverge, since each term is bounded.

It is useful to give an approximate expression of the correlation function in the r space that corresponds to the expressions we derived in the k space. Let us therefore assume that the correlation function in k space is approximately given by

$$C(\boldsymbol{k}) \propto \frac{1}{k^2 + \xi^{-2}}. \tag{5.142}$$

The inverse Fourier transform of this expression is given by

$$C(\boldsymbol{r}) \propto \int \mathrm{d}\boldsymbol{k}\, \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}}{k^2 + \xi^{-2}}. \tag{5.143}$$

In order to evaluate this expression, let us consider the identity

$$\frac{1}{x} = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-tx}. \tag{5.144}$$

We obtain

$$C(\boldsymbol{r}) \propto \int \mathrm{d}\boldsymbol{k} \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-t(k^2+\xi^{-2})+\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{5.145}$$

We can now evaluate the Gaussian integral over the ks and obtain

$$C(\boldsymbol{r}) \propto \int_0^\infty \mathrm{d}t\, t^{-d/2}\, \mathrm{e}^{-t\xi^{-2}-r^2/(4t)}. \tag{5.146}$$

Let us switch the integration variable to w, defined by

$$w = t\, r^{-2}. \tag{5.147}$$

We obtain

$$C(\boldsymbol{r}) \propto r^{2-d} \int_0^\infty \mathrm{d}w\, w^{-d/2} \exp\left[-w\left(\frac{r}{\xi}\right)^2 - \frac{1}{4w}\right]. \tag{5.148}$$

If $d > 2$ and $|\boldsymbol{r}| \ll \xi$, we can neglect the r-dependence in the integral, and we obtain $C(\boldsymbol{r}) \sim r^{2-d}$. On the other hand, if $r \gg \xi$, we can evaluate the integral by the saddle-point method. The saddle point lies at $w = w_c = \xi/2r$. We thus obtain

$$C(\boldsymbol{r}) \sim \mathrm{e}^{-r/\xi}. \tag{5.149}$$

These two behaviors can be summarized by the interpolating expression

$$C(\boldsymbol{r}) \propto \frac{\mathrm{e}^{-r/\xi}}{r^{d-2}}, \tag{5.150}$$

which is exact (for this model) in $d = 3$.

5.15 The Landau Theory

The mean-field theory allows us to obtain an approximate expression for the phase diagram of the system under study. If, however, we want to aim for the more limited goal of describing the system's behavior close to the critical point, we can resort to a theory that is formally simpler and more easily generalizable—the Landau theory.

In mean-field theory we look for the minimum of an effective free energy, defined by

$$\mathcal{F}(h, T; m) = N\left[f(T; m) - hm\right], \tag{5.151}$$

expressed as a function of the magnetization per spin, m. The trial free energy per spin, expressed as a function of m, has the form

$$\begin{aligned} f(T; m) &= -\frac{\zeta J}{2} m^2 - Ts(m) \\ &= -\frac{\zeta J}{2} m^2 \\ &\quad + k_B T \left[\left(\frac{1-m}{2} \right) \ln \left(\frac{1-m}{2} \right) + \left(\frac{1+m}{2} \right) \ln \left(\frac{1+m}{2} \right) \right]. \end{aligned} \tag{5.152}$$

Close to the critical point, the equilibrium value of the order parameter m is small, and we can therefore expand this function in a power series:

$$f(T; m) = f(T; 0) + \frac{1}{2} a(T) m^2 + \frac{1}{4} b(T) m^4 + \cdots. \tag{5.153}$$

Only the even powers of m appear in this expansion, because the Hamiltonian is symmetrical with respect to the transformation $m \longrightarrow -m$. It is obviously possible to calculate the coefficients $a(T)$, $b(T)$, and so on explicitly. For the Ising model on a square lattice, we have, for example,

$$\begin{aligned} f(T; 0) &= k_B T \ln 2, \\ a(T) &= k_B T - \zeta J \propto (T - T_c), \\ b(T) &= k_B T/3. \end{aligned} \tag{5.154}$$

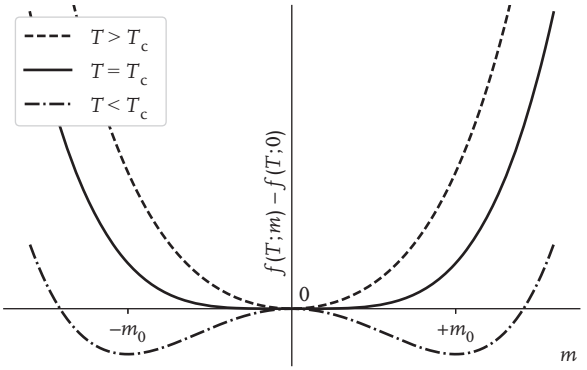


FIGURE 5.14. Landau free energy per spin, $f(T;m)$, as a function of the magnetization m per spin near the critical temperature.

We see that the coefficient a changes sign at the critical temperature, while the coefficient b remains positive. In effect, when we look for the minimum of $\mathcal{F}$ with respect to m, we expect to find a vanishing solution for $T > T_c$ and two opposite minima at $\pm m_0$, where $m_0 = \sqrt{|a|/b}$, for $T < T_c$. The behavior of $f(T;m)$ for small values of m, and for temperatures close to T_c, is shown in figure 5.14.

Staying close to T_c we can set

$$\begin{aligned} &a(T) \simeq a'\,(T - T_c), \quad a' > 0, \\ &b(T) \simeq b(T_c) = b > 0. \end{aligned} \tag{5.155}$$

Then the equation for the minimum of f takes the form

$$a'\,(T - T_c)m + b\,m^3 = h. \tag{5.156}$$

For $T > T_c$ we obtain

$$m \simeq \frac{1}{a'(T - T_c)} h = \chi\, h, \tag{5.157}$$

which confirms the Curie-Weiss law $\chi \propto (T - T_c)^{-1}$.

On the other hand, for $T < T_c$, we have *three* solutions. One is of the form

$$m \simeq \frac{1}{a'\,(T - T_c)} h, \tag{5.158}$$

which correponds to a local *maximum* of $f(T;m)$, as is clear from the figure. The other two, $m_\pm$, correspond to local minima and are given by the equation

$$b\,m_\pm^2 = a'\,|T - T_c|\,, \tag{5.159}$$

approximate up to terms that vanish as $h \to 0$. We thus obtain the behavior of the spontaneous magnetization per spin $m_0(T)$ close to T_c:

$$m_0(T) = \lim_{h \to 0^+} m_+(T,h) = \left(\frac{a'\,|T - T_c|}{b} \right)^{1/2} \propto |T - T_c|^{1/2}\,. \tag{5.160}$$

It is also interesting to evaluate the behavior of m exactly at the critical temperature. We have

$$b\,m^3 = h, \tag{5.161}$$

and we obtain therefore $m \propto h^{1/3}$.

One can also evaluate the specific heat. In fact, since the entropy per particle is given by

$$s = -\left.\frac{\partial f}{\partial T}\right)_h, \tag{5.162}$$

one obtains for the specific heat per particle

$$\frac{C_h}{NT} = -\left.\frac{\partial^2 f}{\partial T^2}\right)_h = -f''(T;0) + \begin{cases} 0, & \text{for } T > T_c, \\ a'^2/(2b), & \text{for } T < T_c. \end{cases} \tag{5.163}$$

The specific heat is therefore discontinuous at the transition.

All the results for the critical behavior of thermodynamic quantities derive from the hypothesis that the thermodynamic value of the order parameter can be obtained by means of a variational principle, where a function $f(T;m)$ appears that is analytic and symmetrical with respect to a certain transformation group of m. In our case, this group is composed of the identity and the inversion $m \longrightarrow -m$.

One can consider more general cases, in which the symmetry group is more complicated. In such cases, in the expansion of $f(T;m)$, only the invariants that can be built starting from m can appear. Let us suppose for example that the order parameter is a vector $\boldsymbol{m} = (m_a)$ in n dimensions and that it transforms according to the rotation group in n dimensions, $\mathsf{O}(n)$. The only invariant we can build is $m^2 = \sum_\alpha m_\alpha^2$. Therefore, $f(T;m)$ is in fact only a function of m^2, and its expansion into a Taylor series has the same form as that in the Ising model. The results we obtained are therefore still essentially valid. There are, however, some effects related to the vectorial nature of the order parameter. In particular, we can distinguish between the effect of a field applied in the direction of the order parameter $\boldsymbol{m}$ and that of one normal to it. For $T < T_c$, the susceptibility to a field parallel to $\boldsymbol{m}$ is finite for $h \to 0$, while the susceptibility to a field normal to it diverges as h^{-1}.

Exercise 5.13 Prove the above assertion by applying the Landau theory to a system with vector order parameter m.

On the other hand, in more complicated situations, there can be cubic invariants, or invariants proportional to higher odd powers of m. When this happens, the Landau theory predicts that the transition becomes discontinuous. In effect, exactly at the critical temperature, f has to admit a minimum at $m = 0$, but this is impossible if there is a term proportional to m^3. Cubic invariants of this type are always present at the fluid–solid transition-therefore, the order parameter of the fluid-solid transition cannot vanish continuously (at least within the context of mean-field theory).

The analyticity of f is due to the very nature of the mean-field approximation—the effect of the rest of the sample on a degree of freedom (or on a small number of degrees of

freedom) is represented by effective interaction parameters, such as the effective magnetic field λ. The trial free energy is then obtained by evaluating the partition sum over these few degrees of freedom—it is obvious that this procedure cannot produce nonanalytic dependencies in the order parameter.

It is possible, in principle, to improve the mean-field approximation by considering a larger, but finite, number of degrees of freedom—this improvement, however, cannot modify those results of the mean-field theory that depend only on the analyticity of the trial free energy.

Exercise 5.14 (Tricritical point) Consider a system whose free energy density is given by

$$f(T, h, b; m) = \text{const.} + \frac{1}{2}am^2 + \frac{1}{4}bm^4 + \frac{1}{6}cm^6 - hm, \tag{5.164}$$

where m is a scalar and $c > 0$ is fixed. Set $h = 0$ at first.

1. Evaluate for fixed $b > 0$ the extrema of the free energy density as a function of a, and discuss their stability. Find the location of the critical point and evaluate the magnitude of the order parameter in the broken-symmetry phase.

2. Identify for fixed $b < 0$ the minima of the free energy and show that there is a point in which the symmetry breaks spontaneously. Discuss the nature of the transition. Sketch the phase diagram in the (a, b) plane, distinguishing between continuous and discontinuous transitions.

3. Show that for $h = 0$ there are surfaces in the (a, b, h) space where a discontinuous transition between two different broken-symmetry phases takes place, and identify what bounds these surfaces.

4. The point $(a, b, h) = (0, 0, 0)$ is called the **tricritical point**. Evaluate the thermodynamic critical exponents approaching the tricritical point along the line $b = 0$.

5. Show that for b small but positive, as one approaches the critical temperature, one should first observe a behavior dictated by the tricritical point, and eventually those characteristic of the critical one.

5.16 Critical Exponents

We have obtained the behavior of various thermodynamic quantities in the neighborhood of the critical point. These behaviors are described by power laws, whose exponents are called **critical exponents**. The results we obtained are summarized in table 5.1. These values of the critical exponents follow from the analyticity hypothesis and are called the classical values. Given the generality of the hypotheses underlying the Landau theory, it was a surprise to see that the experimentally observed values of the critical exponents are not the classical ones. Actually, physical systems can be grouped into broad classes, called **universality classes**: within each class the exponents have the same value, and it is in fact

Table 5.1 Classical critical exponents.

Quantity	Behavior	Region	Exponent
C_h	$\lvert T - T_c\rvert^{-\alpha}$	$h = 0$	$\alpha = 0$ (discont.)
χ_T	$\lvert T - T_c\rvert^{-\gamma}$	$h = 0$	$\gamma = 1$
m	$\lvert T - T_c\rvert^{\beta}$	$h = 0,\ T < T_c$	$\beta = 1/2$
	$\lvert h\rvert^{1/\delta}$	$T = T_c$	$\delta = 3$
ξ	$\lvert T - T_c\rvert^{-\nu}$	$h = 0$	$\nu = 1/2$
$C(\boldsymbol{k})$	$\lvert\boldsymbol{k}\rvert^{-2+\eta}$	$h = 0,\ T = T_c$	$\eta = 0$

possible to map the critical behavior of one system into that of another by means of simple transformations.

In table 5.1, we have also introduced the exponent η that describes the behavior of the correlation function in Fourier space for $T = T_c$ and $h = 0$:

$$C(\boldsymbol{k}) \propto \lvert\boldsymbol{k}\rvert^{-2+\eta}. \tag{5.165}$$

This exponent vanishes in mean-field theory.

Two challenges remain:

1. Identify the reason for the lack of success of the Landau theory.
2. Identify the reason for the existence of the universality classes, and if possible, provide a method for calculating the critical behavior within each class.

Observe that the hypothesis that the order parameter has a well-specified value, obtained from the variational principle on the free energy, corresponds to the hypothesis that it is possible to disregard its fluctuations. We can expect that the theory is not reliable when the fluctuations of the order parameter dominate its mean value. This leads to a useful criterion (the **Ginzburg criterion**) for the range of validity of the Landau theory. In order to make this criterion explicit, however, it is necessary to describe the behavior of the critical fluctuations.

5.17 The Einstein Theory of Fluctuations

In order to understand the critical behavior, it is necessary to treat the fluctuations of thermodynamic quantities in small, but macroscopic, systems. This theory of fluctuations is essentially due to Einstein.

In order to clarify our ideas, let us consider a system described by the Ising model, at a given value of the temperature T and magnetic field h. Within this system, let us focus our attention on a small but macroscopic region S. We want to evaluate the probability that the energy and the magnetizations of the spins contained in this region have respectively the values E_S and M_S.

To evaluate this probability, let us consider those microstates ν_S of the subsystem S such that the internal energy of the subsystem is equal to E_S and its magnetization is equal

to M_S. The probability that the subsystem is in any microstate of this kind is proportional to the Boltzmann factor $\exp[-(E_S - hM_S)/k_B T]$, with the given values of T and h. Thus, the probability that the subsystem S has these values of internal energy and magnetization is obtained by summing this probability over all the microstates satisfying the condition just mentioned:

$$p(E_S, M_S) = \frac{1}{Z} \sum_{\nu_S} \delta\left[M_S - M(\nu_S)\right] \delta\left[E_S - E(\nu_S)\right] \times \exp\left(-\frac{E_S - hM_S}{k_B T}\right). \tag{5.166}$$

In this expression, Z is the partition function, which imposes the normalization condition on the probability distribution of the microstates of S, and the delta functions impose the condition on the values of E and M.

By the fundamental postulate, the sum over all microscopic states gives the exponential of the entropy, expressed as a function of the internal energy and the magnetization:

$$\sum_{\nu_S} \delta\left[M_S - M(\nu_S)\right] \delta\left[E_S - E(\nu_S)\right] = \exp\left[\frac{S_S(E_S, M_S)}{k_B}\right]. \tag{5.167}$$

Therefore

$$p(E_S, M_S) = \frac{1}{Z} \exp\left[-\frac{E_S - TS_S(E_S, M_S) - hM_S}{k_B T}\right] = \exp\left[-\frac{\Delta\mathcal{F}(E_S, M_S; T, h)}{k_B T}\right], \tag{5.168}$$

where

$$\Delta\mathcal{F}(E_S, M_S; T, h) = E_S - TS_S(E_S, M_S) - hM_S - F(T, h) \tag{5.169}$$

is the **availability**, and $F(T, h) = -k_B T \ln Z$ is the Helmholtz free energy, expressed as a function of the temperature and magnetic field. Note that the availability vanishes for the equilibrium values of energy and magnetization by virtue of the variational principle of the Legendre transform. The availability is often called *free energy difference*—let us note, however, that unlike the actual free energy, it is a function both of intensive variables (in our case, the magnetic field and the temperature) and of their conjugate fluctuating extensive variables.

Let us observe that we have obtained an apparently nonnormalized probability distribution. Indeed, the quantities that appear explicitly in eq. (5.169) are extensive, and are defined up to subextensive corrections. One can easily check that, since the normalization factor behaves like a power of the subsystem size, imposing the normalization also leads to subextensive corrections, which are therefore negligible.

If, as usually occurs, entropy is a regular function of its arguments close to the equilibrium values, then it is possible to deduce a certain number of important relations for the

fluctuations. Let us first consider a system enclosed by nonadiabatic walls, in which the internal energy can fluctuate. We then have

$$\Delta\mathcal{F}(E;T) \simeq -\frac{T}{2}\left.\frac{\partial^2 S}{\partial E^2}\right|_{E=E^{\mathrm{eq}}} \Delta E^2, \tag{5.170}$$

where $\Delta E = E - E^{\mathrm{eq}}$, and E^{eq} is the equilibrium value of the energy. We then obtain

$$\begin{aligned}\langle \Delta E^2 \rangle &= \frac{1}{\mathcal{Z}} \int \mathrm{d}\Delta E\, \Delta E^2 \, \exp\left[\frac{1}{2k_{\mathrm{B}}} \frac{\partial^2 S}{\partial E^2}\right)_V \Delta E^2\Bigg] = -k_{\mathrm{B}} \frac{\partial^2 S}{\partial E^2}\Bigg)_V^{-1} \\ &= -k_{\mathrm{B}} \frac{\partial (1/T)}{\partial E}\Bigg)_V^{-1} = k_{\mathrm{B}} T^2 \frac{\partial E}{\partial T}\Bigg)_V = k_{\mathrm{B}} T^2 C_V.\end{aligned} \tag{5.171}$$

This result obviously coincides with the one obtained for the fluctuations in the canonical ensemble.

Let us now focus our attention on a collection (X_i), $i = 1, 2, \ldots, r$, of extensive quantities. Let us denote the respective equilibrium values with X_i^{eq}. The first nonvanishing term of the Taylor expansion of availability is the second one:

$$\frac{\Delta\mathcal{F}}{k_{\mathrm{B}} T} \simeq \frac{1}{2k_{\mathrm{B}}} \sum_{ij} \frac{\partial^2 S}{\partial X_i\, \partial X_j} \Delta X_i\, \Delta X_j, \tag{5.172}$$

where $\Delta X_i = X_i - X_i^{\mathrm{eq}}$. The linear terms of the expansion vanish due to the variational principle. Because of the concavity of the entropy, this quadratic form is positive-semidefinite. More particularly, if it is positive-definite, and if it is possible to disregard further terms of the Taylor expansion, we deduce from eq. (5.172) that the distribution of the fluctuations of the (X_i) is multivariate Gaussian, defined by the matrix $\mathsf{K} = (K_{ij}) = -(\partial^2 S/\partial X_i \partial X_j)/k_{\mathrm{B}}$. The correlation matrix C is the inverse of K (cf. sec. A.7):

$$C_{ij}^{-1} = -\frac{1}{k_{\mathrm{B}}} \frac{\partial^2 S}{\partial X_i\, \partial X_j} = \frac{1}{k_{\mathrm{B}}} \frac{\partial (f_i/T)}{\partial X_j^{\mathrm{eq}}}\Bigg)_X . \tag{5.173}$$

We have made explicit the fact that the derivative is taken by keeping the average values of the other extensive variables constant. This result coincides with what can be obtained by considering the corresponding generalized ensemble:

$$C_{ij} = \langle \Delta X_i\, \Delta X_j \rangle = \frac{\partial^2 \ln Z}{\partial (f_i/k_{\mathrm{B}}T)\, \partial (f_j/k_{\mathrm{B}}T)} = k_{\mathrm{B}} \frac{\partial X_i^{\mathrm{eq}}}{\partial (f_j/T)}\Bigg)_{f/T} . \tag{5.174}$$

In this case, the derivative is taken by keeping the values of the ratios f/T constant. Recall that the force conjugate to the internal energy is $1/T$. The expressions of $\mathsf{C} = (C_{ij})$ are yielded by eqs. (5.173) and (5.174) are equivalent, since in effect one has

$$\sum_{\ell} \frac{\partial X_i^{\mathrm{eq}}}{\partial (f_\ell/T)}\Bigg)_{f/T} \frac{\partial (f_\ell/T)}{\partial X_j^{\mathrm{eq}}}\Bigg)_X = \delta_{ij}. \tag{5.175}$$

This approach to the statistics of fluctuations is not limited to just the usual thermodynamic quantities. For example, it can be generalized to the study of the spatial distribution of the magnetization density. For example, we can evaluate the probability that a given

Fourier component of the magnetization density has a given value ϕ_k. The component of the magnetization density is an observable, because it is expressed by

$$\phi_k = \sum_j \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j}\sigma_j. \tag{5.176}$$

Its probability is proportional to $\exp\left[-\Delta\mathcal{F}(\phi_k)/k_\mathrm{B}T\right]$, where, for small fluctuations, one has

$$\frac{\Delta\mathcal{F}(\phi_k)}{k_\mathrm{B}T} \simeq \frac{1}{2N}\frac{|\phi_k|^2}{C(\boldsymbol{k})}. \tag{5.177}$$

In this expression, $C(\boldsymbol{k})$ is the Fourier transform of the correlation function:

$$C(\boldsymbol{k}) = \frac{1}{N}\left\langle|\phi_k|^2\right\rangle = \sum_j \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_{ij}}\left\langle\sigma_i\sigma_j\right\rangle_\mathrm{c}. \tag{5.178}$$

For small values of $|\boldsymbol{k}|$, one can approximate $C(\boldsymbol{k})$ by means of its Ornstein-Zernike expression:

$$C(\boldsymbol{k}) = \frac{k_\mathrm{B}T\,\chi}{1+k^2\xi^{-2}}, \tag{5.179}$$

where ξ is the coherence length, and χ the susceptibility per spin. The expressions of χ and ξ are not necessarily given correctly by mean-field theory. In order for this expression to be valid, it is sufficient that $C^{-1}(\boldsymbol{k})$ can be expanded in a Taylor series as a function of k^2—in other words, that the correlations be of finite range.

We have so far considered the fluctuations of extensive quantities in a finite region of the system, which are in contact with a much larger system, and therefore characterized by well-defined values of the corresponding intensive quantities (generalized forces and temperature). If we consider an isolated system, we can ask ourselves whether it is possible to observe analogous fluctuations in the intensive quantities. This problem is still the subject of discussion. One can formally associate corresponding fluctuations of the intensive variables with the fluctuations of the extensive variables by means of the equations of state. This procedure is fairly arbitrary, however, and, above all, does not allow one to obtain further physical insights.

5.18 Ginzburg Criterion

Let us now consider the fluctuations of the magnetization as we approach the critical temperature. We can estimate the availability by the mean-field theory (or, equivalently, by the Landau theory):

$$\Delta\mathcal{F} \simeq \frac{N}{2}a'(T-T_\mathrm{c})m^2 + \frac{Nb}{4}m^4. \tag{5.180}$$

As long as $T > T_\mathrm{c}$, the first term dominates the second one when $N \to \infty$. The fluctuations of m are therefore Gaussian and $\mathrm{O}(N^{-1/2})$. The variance diverges as we get closer to the

critical temperature, and exactly at T_c, the fluctuations are no longer Gaussian:

$$p(m) \propto \exp\left(-\frac{Nb}{4k_B T} m^4\right). \tag{5.181}$$

Close to the critical temperature, the fluctuations become very intense and are spatially correlated over very large distances—we have in fact seen that the coherence length diverges for $T \to T_c$. It is therefore essential to understand when it is possible to neglect the fluctuations, and therefore where the predictions of mean-field theory can be considered valid, and where they become unreliable.

In order to estimate the relative importance of the fluctuations, we employ the **Ginzburg criterion** [Ginz60]. The idea is to compare the mean value of the order parameter as predicted by mean-field theory with the fluctuations predicted by the Ornstein-Zernicke theory.

We can extend the Einstein theory of fluctuations to cover non-uniform configurations of the order parameter m. For slowly varying configurations, the availability $\Delta\mathcal{F}$ presents a term proportional to $|\nabla m|^2$. By suitably rescaling m we obtain

$$\frac{\Delta\mathcal{F}}{k_B T} \simeq \int \mathrm{d}\boldsymbol{r} \left[\frac{a}{2} m^2 + \frac{b}{4} m^4 + \frac{1}{2} |\nabla m|^2\right]. \tag{5.182}$$

Then the coherence length is given by $\xi^{-2} = |a|$, and the spontaneous magnetization is given by $m_0(T)^2 = |a|/b$. The correlation function $C(\boldsymbol{r})$ in $\boldsymbol{r}$ space has the form (see eq. (5.150))

$$\frac{C(\boldsymbol{r})}{k_B T} \simeq \frac{\mathrm{e}^{-r/\xi}}{r^{d-2}}. \tag{5.183}$$

As a measure of the fluctuation intensity, we can take the value of $C(\boldsymbol{r})$ when $|\boldsymbol{r}| = \xi$. We thus obtain

$$\left\langle \Delta m^2 \right\rangle \simeq \xi^{-(d-2)} \simeq |a|^{(d-2)/2}. \tag{5.184}$$

This must be compared with $m_0(T)^2$. We thus obtain that mean-field theory is reliable as long as $|a|^{(d-4)/2} < 1/b$. The left-hand side diverges when $T \to T_c$ for $d < 4$, and therefore the bound is broken for $|a| < b^{2/(d-4)}$. This is the Ginzburg criterion.

The criterion can be made more quantitative by introducing a systematic approximation scheme that reproduces mean-field theory at the lowest order, and by comparing it with the first correction. See, for example, the book by Amit and Martín-Mayor [Amit05, sec. 6.4].

Either way, we obtain that, if $d > 4$, mean-field theory describes accurately the behavior of the system up to the critical temperature, and that, for $d < 4$, there is a range of temperatures around T_c where we should expect deviations from the classical behavior.

5.19 Limit $n \to \infty$

In this section, we consider a system with an n-component vector order parameter in the limit $n \to \infty$, and we show that it can be solved exactly, and that the corresponding critical exponents are not classical.

Let us therefore consider the system defined by the Hamiltonian

$$\mathcal{H} = -\frac{1}{2} \sum_{ij} K_{ij} \boldsymbol{\phi}_i \cdot \boldsymbol{\phi}_j, \tag{5.185}$$

where the interaction matrix (K_{ij}) is defined by

$$K_{ij} = \begin{cases} K_0 = J/k_{\mathrm{B}}T, & \text{if } i \text{ and } j \text{ are nearest neighbors,} \\ 0, & \text{otherwise,} \end{cases} \tag{5.186}$$

and where the $\boldsymbol{\phi}$ are n-component vectors $(\phi_i = (\phi_i^\alpha), \alpha = 1, \ldots, n)$ of magnitude $\sqrt{n}$:

$$\sum_{\alpha=1}^{n} \left(\phi_i^\alpha\right)^2 = n. \tag{5.187}$$

We assume that the system is subject to an external applied field $\boldsymbol{h}$ acting in the direction 1. Then its partition function is given by

$$Z(h) = \int \prod_i \left[\prod_\alpha \mathrm{d}\phi_i^\alpha \, \delta\left(\sum_\beta (\phi_i^\beta)^2 - n \right) \right] \exp\left(\frac{1}{2} \sum_{ij} K_{ij} \sum_\alpha \phi_i^\alpha \phi_j^\alpha + h \sum_i \phi_i^1 \right). \tag{5.188}$$

In order to impose the constraint (5.187), we take advantage of the identity (D.41):

$$\delta(x) = \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}\lambda}{2\pi\,\mathrm{i}} \, \mathrm{e}^{-\lambda x}. \tag{5.189}$$

We have therefore

$$\delta\left(\sum_\alpha (\phi_i^\alpha)^2 - n \right) = \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}\lambda_i}{2\pi\,\mathrm{i}} \, \exp\left[-\lambda_i \left(\sum_\alpha (\phi_i^\alpha)^2 - n \right) \right]. \tag{5.190}$$

In this manner, we can express the partition function (up to constant factors) in the following form:

$$\begin{aligned} Z(h) = \int \prod_i \left(\mathrm{d}\lambda_i \prod_\alpha \mathrm{d}\phi_i^\alpha \right) \\ \times \exp\left[-\frac{1}{2} \sum_\alpha \sum_{ij} \left(2\lambda_i \delta_{ij} - K_{ij} \right) \phi_i^\alpha \phi_j^\alpha + n \sum_i \lambda_i + h \sum_i \phi_i^1 \right]. \end{aligned} \tag{5.191}$$

Since the ϕ_i^α's appear quadratically in this expression, it is possible to evaluate the corresponding functional integral exactly, by exploiting the relation (A.46). Actually, it is more convenient to integrate over all components of $\boldsymbol{\phi}$, except for the first one ϕ_i^1, which we shall denote by σ_i. The integral converges if the matrix $\mathsf{A} = \left(2\lambda_i\delta_{ij} - K_{ij}\right)$ is positive definite. One thus obtains

$$Z(h) = \prod_i \left(\mathrm{d}\lambda_i\, \mathrm{d}\sigma_i\right) \exp\left[\frac{1}{2}\sum_{ij}\left(K_{ij} - 2\lambda_i\delta_{ij}\right)\sigma_i\sigma_j + h\sum_i \sigma_i + n\sum_i \lambda_i - \frac{n-1}{2}\,\mathrm{Tr}\ln\left(2\lambda_i\delta_{ij} - K_{ij}\right)\right]. \tag{5.192}$$

The term in Tr ln is the logarithm of the determinant arising from the Gaussian integral (see eq. (A.46)). This term can be evaluated more easily by shifting to the Fourier transforms. Let us introduce the Fourier transform of K_{ij}:

$$K(\boldsymbol{k}) = \sum_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\left(\boldsymbol{r}_j - \boldsymbol{r}_i\right)} K_{ij}. \tag{5.193}$$

For a d-dimensional simple cubic lattice with lattice constant equal to a_0, one has (as we have seen in sec. 5.14) $K(\boldsymbol{k}) = 2K_0\sum_{\alpha=1}^{d}\cos(k_\alpha a_0)$. We set, for convenience,

$$K(\boldsymbol{k}) = K_0\,\omega(\boldsymbol{k}), \tag{5.194}$$

where, for $ka_0 \ll 1$, one has $\omega(\boldsymbol{k}) \simeq 2d - k^2a_0^2$.

In the limit $n \to \infty$, we can set $n - 1 \simeq n$ and evaluate the remaining integrals with the saddle point method (see sec. E). We look for a translation-invariant saddle point:

$$\sigma_i = \sigma, \qquad \lambda_i = \lambda, \qquad \forall i. \tag{5.195}$$

With these settings, the logarithm $\Delta(\lambda)$ of the determinant of the matrix A can be expressed as follows:

$$\Delta(\lambda) = \mathrm{Tr}\ln\left(2\lambda\delta_{ij} - K_{ij}\right) = \sum_{\boldsymbol{k}}\ln\left[2\lambda - K_0\omega(\boldsymbol{k})\right] = \frac{Na_0^d}{(2\pi)^d}\int_{\mathrm{B.Z.}}\mathrm{d}\boldsymbol{k}\,\ln\left[2\lambda - K_0\omega(\boldsymbol{k})\right], \tag{5.196}$$

where the integral over $\boldsymbol{k}$ is extended to the first Brillouin zone.

The saddle-point equations are then the following:

$$\left[2\lambda - K_0\omega(0)\right]\sigma = h, \tag{5.197a}$$

$$-\sigma^2 + n = \frac{n}{2N}\frac{\mathrm{d}}{\mathrm{d}\lambda}\Delta(\lambda). \tag{5.197b}$$

One can easily see that

$$\frac{1}{2N}\frac{\mathrm{d}}{\mathrm{d}\lambda}\Delta(\lambda) = \frac{a_0^d}{(2\pi)^d}\int_{\mathrm{B.Z.}}\frac{\mathrm{d}\boldsymbol{k}}{2\lambda - K_0\omega(\boldsymbol{k})}. \tag{5.198}$$

For $h = 0$, the first saddle point equation (5.197a) implies that $\sigma = 0$ or $2\lambda = K_0\omega(0)$. Let us consider the first case, which, as we will see, holds above the critical temperature. The second equation (5.197b) becomes a consistency equation for λ, of the form

$$1 = f(\lambda), \tag{5.199}$$

where

$$f(\lambda) = \frac{a_0^d}{(2\pi)^d} \int_{\text{B.Z.}} \frac{\mathrm{d}\boldsymbol{k}}{2\lambda - K_0\omega(\boldsymbol{k})}. \tag{5.200}$$

Let us discuss this expression. In order for the integral over the $\boldsymbol{\phi}$ that led us to this expression to be convergent, all the eigenvalues of the matrix $\left(2\lambda_i\delta_{ij} - K_{ij}\right)$ must be positive. This implies that $2\lambda > K_0\omega(0)$—in other words, $2\lambda/K_0 > 2d$. Let us consider the behavior of the quantity $K_0 f(\lambda)$ as a function of λ/K_0. It increases as λ/K_0 decreases, and therefore its maximum value is reached when $2\lambda/K_0 = \omega(0)$.[1] Let us introduce the new variable $r = [2\lambda/K_0 - \omega(0)]/a_0^2$, and calculate the value of $K_0 f(\lambda)$ for small values of r:

$$K_0 f(\lambda) \propto \int_{\text{B.Z.}} \mathrm{d}\boldsymbol{k}\, \frac{1}{r + k^2} + \int_{\text{B.Z.}} \mathrm{d}\boldsymbol{k} \left[\frac{a_0^2}{2\lambda - K_0\omega(\boldsymbol{k})} - \frac{1}{r + k^2} \right]. \tag{5.201}$$

The second term of this expression is convergent and analytic, even for $r = 0$. Let us discuss the first term. One has

$$\mathcal{I} = \int_{\text{B.Z.}} \mathrm{d}\boldsymbol{k}\, \frac{1}{r + k^2} = \text{const.} + \mathrm{O}\left(r^{(d-2)/2}\right). \tag{5.202}$$

We can distinguish between two cases:

1. $d \leq 2$: $\lim_{r\to 0} \mathcal{I} = \infty$.
2. $d > 2$: $\mathcal{I} = \text{const.} + \mathrm{O}(r^s)$, where the exponent $s = (d-2)/2$ is smaller than 1 for $2 < d < 4$ and larger than 1 otherwise.

In the first case, for each temperature value, eq. (5.197b) admits a solution with a finite λ and $\sigma = 0$. In other words, there is no spontaneous symmetry breaking. In the second case, it is possible to define a critical value for K_c by means of the equation

$$K_c^2\omega(0) = \frac{2a_0^d}{(2\pi)^d} \int_{\text{B.Z.}} \frac{\mathrm{d}\boldsymbol{k}}{\omega(0) - \omega(\boldsymbol{k})}. \tag{5.203}$$

For $K > K_c$, eq. (5.197b) no longer admits solutions with $\sigma = 0$. We can therefore assume that $\sigma > 0$, which, because of eq. (5.197a), implies that for $h = 0$, $2\lambda = K_0\omega(0)$. Equation (5.197b) therefore becomes an equation for σ. We can easily see that $\sigma \sim n^{1/2}$, and we can therefore rescale it by defining $\sigma^2 = nm^2$. With a little algebra, one thus obtains

$$m^2 = 1 - K_0/K_c, \tag{5.204}$$

[1] The analogy with the treatment of the Einstein condensation should be evident.

from which we can read the value of exponent β:

$$\beta = \frac{1}{2}. \tag{5.205}$$

In spite of this result, the exponents of this model do not have the classical values, at least for $2 < d < 4$. It is actually easy to see that in this case one has $2\lambda - K_c \propto r \propto |K_c - K_0|^{2/(d-2)}$ for $K_0 \to K_c$. Now from eq. (5.197a), we can obtain the expression per susceptibility per spin χ above the critical temperature (in other words, for $K_0 < K_c$, $h = 0$):

$$\chi = \left.\frac{\partial \sigma}{\partial h}\right|_{h=0} = \frac{1}{2\lambda - K_0 \omega(0)}. \tag{5.206}$$

We thus obtain the critical behavior of the susceptibility and therefore the exponent γ:

$$\chi \propto |K_c - K_0|^{-\gamma}; \qquad \gamma = \begin{cases} 2/(d-2), & \text{if } 2 < d < 4, \\ 1, & \text{if } d > 4. \end{cases} \tag{5.207}$$

(A more detailed analysis shows that for $d = 4$, logarithmic corrections are present.) From the expression of the correlation function, which is proportional to $1/(r + k^2)$, one also obtains the behavior of the coherence length ξ:

$$\xi = r^{-1/2} \propto |K_c - K_0|^{-\nu}; \qquad \nu = \begin{cases} 1/(d-2), & \text{if } 2 < d < 4, \\ 1/2, & \text{if } d > 4. \end{cases} \tag{5.208}$$

The advantage of this model is that it exhibits nonclassical behavior that can be calculated exactly, without resorting to the renormalization methods that we shall study in the next chapter. The $n \to \infty$ limit is a powerful tool for the study of complex systems.

Exercise 5.15 (Spherical model) Consider a model defined by the Hamiltonian (5.38) defined on a lattice in d dimensions. Here the variables σ_i do not satisfy $\sigma_i^2 = 1$, $\forall i$, but the much weaker constraint

$$\sum_i \sigma_i^2 = N,$$

where N is the number of points in the lattice. Thus $\sigma = (\sigma_i)$ belongs to a sphere in N dimensions, which justifies the name "spherical model." Show that this model can be solved exactly, and that its behavior corresponds to that of the $n \to \infty$ limit.

5.20 Universality and Scaling

It has been possible to evaluate the critical exponents for a few nontrivial, exactly solvable models. We also have numerical or experimental evidence for their values in a number of real systems. It turns out that these values are different from the classical ones for the most familiar physical systems, such as the liquid-gas critical point or different kinds of

Table 5.2 Critical exponents of different physical systems.

	Gas-Liquid [Rowl82a]	Binary Mixtures [Rowl82b]	β-brass [AlsN76]	^{4}He [Ahle80]	Ferromagnets [Kada67]
n	1	1	1	2	3
α	$0.10 \pm .05$	$0.10 \pm .06$		-0.026	$-0.09 - -0.012$
β	$0.32 \pm .01$	$0.32 \pm .02$	$0.305 \pm .005$		$0.32 - 0.36$
γ	$1.24 \pm .05$	$1.25 \pm .05$			$1.33 \pm .015$
δ	$4.8 \pm .02$	$4.5 \pm .5$			4.3 ± 1
η			$0.03 - 0.06$		
ν			$0.65 \pm .02$	$0.672 \pm .001$	$0.69 \pm .02$

Note: The first three columns correspond to systems with a scalar order parameter ($n = 1$): several liquid-vapor system at their critical point, different binary mixtures at their consolution point, and β-brass, a form of copper and tin alloy. The λ-transition ($He^I \longrightarrow He^{II}$) corresponds to a planar order parameter ($n = 2$), while the systems denoted as ferromagnets correspond to ($n = 3$). I have included only exponents that have been actually measured. More specifically, in the case of the λ transition of helium, the order parameter is not accessible, and therefore the exponents β, γ, δ, and η cannot be measured.

ferromagnets. In table 5.2, I report the values of the critical exponents for some physical systems. These values are clearly different from the classical ones. One notices, however, that the exponents of the first three systems are mutually compatible—the exponents seem to depend only on the dimensionality of the order parameter. This conjecture is corroborated by considering the exponents obtained numerically (with different methods) for some statistical mechanics models on a lattice, which are shown in table 5.3. Analogously, the measurable exponents of the λ-transition of helium correspond to those of the XY model (cf. sec. 5.8), while the ferromagnetic or antiferromagnetic transitions characterized by a three-dimensional order parameter are described by the same exponents. These observations have led to the formulation of the universality hypothesis, according to which phase transitions can be sorted into a small number of **universality classes**, characterized by the dimensions of the system and of the order parameter, such that the critical exponents are equal within each class.

One can obtain in fact a stronger relation. Let us consider the equation of state for the order parameter, expressed as a function of h and of

$$t = \frac{T - T_c}{T_c}, \tag{5.209}$$

for two different systems belonging to the same universality class:

$$\begin{aligned} m^{(1)} &= m^{(1)}(t, h); \\ m^{(2)} &= m^{(2)}(t, h). \end{aligned} \tag{5.210}$$

Table 5.3 Best theoretical estimates of the critical exponents for different models as a function of the dimensionality *n* of the order parameter and the dimensionality *d* of space.

	Mean Field	Ising $2d$	Ising $3d$	XY	Heisenberg	$n \to \infty$
(n, d)		$(1, 2)$	$(1, 3)$	$(2, 3)$	$(3, 3)$	$(\infty, 3)$
α	0 (disc.)	0 (log.)	$0.103 \pm .005$	$-0.022 \pm .009$	$-0.133 \pm .018$	-1
β	$1/2$	$1/8$	$0.329 \pm .009$	$0.350 \pm .007$	$0.365 \pm .012$	$1/2$
γ	1	$7/4$	$1.2385 \pm .0025$	$1.322 \pm .005$	$1.400 \pm .006$	2
δ	3	(15)	$4.80 \pm .05$	$4.69 \pm .12$	$4.65 \pm .29$	5
η	0	$1/4$	$0.036 \pm .005$	$0.038 \pm .005$	$0.0375 \pm .005$	0
ν	$1/2$	1	$0.631 \pm .002$	$0.674 \pm .003$	$0.710 \pm .006$	1

Note: The exponent δ of the two-dimensional Ising model is conjectured; the others are obtained from the exact solution. Exponents for the three-dimensional $n \to \infty$ model obtained in sec. 5.19 are also shown. Adapted from [Zinn96, p. 638].

It is then possible to map the two equations of state onto each other by rescaling m, h, and t by an appropriate positive factor:

$$m^{(1)}(t, h) = \lambda_m m^{(2)}(\lambda_t t, \lambda_h h)\,. \tag{5.211}$$

A relation of this kind, on the other hand, must also be valid if one is discussing the *same system*—by arbitrarily choosing one of the scale factors, we can choose the other two so that

$$m(t, h) = \lambda_m m(\lambda_t t, \lambda_h h). \tag{5.212}$$

Let us consider what occurs at $h = 0$, $t < 0$. On the one hand, we obtain

$$m \propto |t|^{\beta}\,, \tag{5.213}$$

and on the other,

$$m(t, h{=}0) = \lambda_m m(\lambda_t t, h{=}0). \tag{5.214}$$

By choosing $\lambda_t = |t|^{-1}$, we obtain

$$\lambda_m = \lambda_t^{-\beta}. \tag{5.215}$$

We now set $t = 0$, $h \neq 0$, and follow the same reasoning to obtain

$$\lambda_m = \lambda_h^{-1/\delta}. \tag{5.216}$$

The magnetization is therefore a general homogeneous function of its arguments:

$$m(\lambda t, \lambda^{\beta\delta} h) = \lambda^{\beta} m(t, h). \tag{5.217}$$

Since the magnetization is the derivative of the free energy with respect to h, an analogous relation also exists for free energy density (or, more exactly, for its singular part):

$$f(\lambda t, \lambda^{\beta\delta} h) = \lambda^x f(t, h). \tag{5.218}$$

We can determine the exponent x from the behavior of specific heat.

Since

$$C_h \propto \frac{\partial^2 f}{\partial t^2}, \tag{5.219}$$

one has

$$C_h(\lambda t, 0) = \lambda^{x-2} C_h(t, 0), \tag{5.220}$$

from which we obtain

$$x = 2 - \alpha. \tag{5.221}$$

The free energy density therefore satisfies the general homogeneity relation:

$$f(\lambda t, \lambda^{\beta\delta} h) = \lambda^{2-\alpha} f(t, h). \tag{5.222}$$

This expression vanishes at the critical point—obviously, this does not mean that all the free energy vanishes at the critical point, but only its singular part.

By taking the derivative of this relation, one can express the exponents of the thermodynamic quantities as a function of only two exponents, α and $\beta\delta$, which appear in the expression of the free energy. By taking its derivative with respect to h, for example, one obtains the expression of the magnetization:

$$\lambda^{\beta\delta} m(\lambda t, \lambda^{\beta\delta} h) = \lambda^{2-\alpha} m(t, h). \tag{5.223}$$

One therefore has

$$\beta = 2 - \alpha - \beta\delta. \tag{5.224}$$

By taking the derivative once more, one obtains the relation for the susceptibility;

$$\lambda^{2\beta\delta} \chi(\lambda t, \lambda^{\beta\delta} h) = \lambda^{2-\alpha} \chi(t, h), \tag{5.225}$$

from which we obtain the exponent γ:

$$-\gamma = 2 - \alpha - 2\beta\delta. \tag{5.226}$$

These two expressions imply a relation between the first three exponents:

$$\alpha + 2\beta + \gamma = 2. \tag{5.227}$$

Relations of this type (known as **scaling laws**) can be written between any three exponents, and follow from the fact that only two exponents appear in the expression of the free energy

exhibiting its homogeneity. They are well satisfied by the experimentally measured exponents, and currently there are few doubts that they are valid for the asymptotic exponents. Another relation of this type, which connects β, γ, and δ, can be obtained by comparing (5.224) and (5.226):

$$\gamma + \beta = \beta\delta. \tag{5.228}$$

It is enlightening to consider the implication of these relations for the coherence length and the correlation function. The lines of reasoning we have followed show that the coherence length, too, must satisfy a general homogeneity relation:

$$\xi(\lambda t, \lambda^{\beta\delta} h) = \lambda^{-\nu} \xi(t, h). \tag{5.229}$$

An analogous relation must also hold for the correlation function, where, however, one has to rescale the wave vector:

$$C(\lambda^{\gamma} k, \lambda t, \lambda^{\beta\delta} h) = \lambda^{z} C(k, t, h). \tag{5.230}$$

In order to determine the exponents γ and z, observe that for $k = 0$, the correlation function is proportional to the susceptibility. We thus obtain

$$z = -\gamma. \tag{5.231}$$

On the other hand, at $t = h = 0$, we must have

$$C(k) \propto k^{-2+\eta}. \tag{5.232}$$

Therefore

$$\frac{\gamma}{\gamma} = 2 - \eta. \tag{5.233}$$

Then, recalling that the coherence length is defined by

$$\xi^2 \propto \frac{1}{C(0)} \left. \frac{\partial C(k)}{\partial k^2} \right|_{k=0}, \tag{5.234}$$

and comparing with the relation for the ξ, we obtain

$$\gamma = \nu. \tag{5.235}$$

We have thus obtained another scaling law:

$$\frac{\gamma}{\nu} = 2 - \eta. \tag{5.236}$$

And lastly, we want to connect the exponent ν to the other thermodynamic variables. We start from the homogeneity relation (5.222), which implies that the singular part of the free-energy density behaves as $|t|^{2-\alpha}$ for $|t| \to 0$. According to the Einstein theory of fluctuations, the typical free energy of an order-parameter fluctuation is of order $k_B T$ (and thus essentially constant) and its extension is of the order of the coherence length ξ. Therefore

the contribution of the order parameter fluctuations to the free-energy density behaves asymptotically as $\xi^{-d} \sim |t|^{dv}$, where d is the dimensionality of the space. We thus obtain the scaling law:

$$2 - \alpha = dv. \tag{5.237}$$

Notice that the classical exponents satisfy all the scaling laws that we have derived so far, except the last one, where the space dimensionality d appears.

Exercise 5.16 Obtain a scaling relation connecting δ, d, and η. Evaluate the exponent δ of the $n \to \infty$ model discussed in sec. 5.19 and check on it the relation.

5.21 Partition Function of the Two-Dimensional Ising Model

In this section, I present the solution of the two-dimensional Ising model in zero magnetic field, as derived by Vdovichenko ([Vdov64] and [Vdov65]). The model was first solved by Lars Onsager in 1944 [Onsa44], by means of a mathematical tour de force. Simpler derivations were found later by Kac and Ward [Kac52], Schulz, Mattis, and Lieb [Schu64], and others. The present solution was inspired by Kac and Ward's solution, but is simpler. It became widely known due to Landau and Lifshitz's treatise on theoretical physics [Land80, pp. 498ff].

We consider a system of $N = L^2$ Ising spins placed on a square lattice. Thus the spin placed at the (k, ℓ) lattice point is denoted by $\sigma_{k\ell}$, and one has $\sigma_{k\ell} = \pm 1$, $k, \ell \in \{1, \ldots, L\}$. The Hamiltonian $\mathcal{H}(\sigma)$ is given by

$$\mathcal{H}(\sigma) = -\sum_{k\ell} \left[J \left(\sigma_{k\ell}\sigma_{k,\ell+1} + \sigma_{k\ell}\sigma_{k+1,\ell} \right) + h\sigma_{k\ell} \right], \tag{5.238}$$

where we have assumed periodic boundary conditions:

$$\sigma_{k+L,\ell} = \sigma_{k,\ell+L} = \sigma_{k\ell}, \forall k, \ell. \tag{5.239}$$

We set $h = 0$ from now on. We have seen in section 5.11, eq. (5.92), that the partition function can be written

$$Z(K) = \left(\frac{2}{1 - t^2} \right)^N \sum_{\mathcal{G}} t^{|\mathcal{G}|}, \tag{5.240}$$

where $K = J/k_B T$,

$$t = \tanh K, \tag{5.241}$$

and the sum runs over all diagrams $\mathcal{G}$ that can be drawn on the lattice, such that (1) each bond appears at most once, and (2) at each vertex only an even number of bonds (zero, two, or four) can meet. In this expression, $|\mathcal{G}|$ is the number of bonds that appear in the diagram $\mathcal{G}$. This expression can be written in the form

$$\mathfrak{S} = \sum_{\mathcal{G}} t^{|\mathcal{G}|} = \sum_r t^r g_r, \tag{5.242}$$

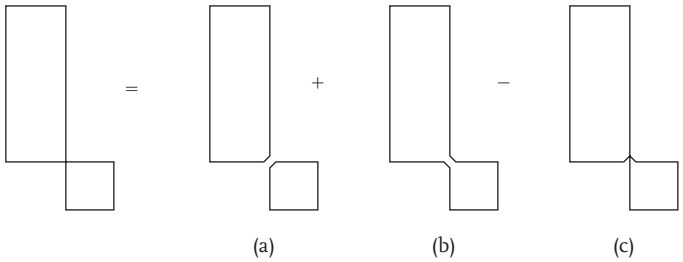


FIGURE 5.15. A diagram with self-intersections can be decomposed in several different ways into loops.

where g_r is the total number of diagrams satisfying the two preceding rules and containing exactly r bonds.

We will now evaluate this expression by transforming it into a sum over loops. The resulting expression will then be evaluated by reducing it to a random-walk problem.

A generic diagram $\mathcal{G}$ can be considered as a collection of **loops**. A loop is the trajectory of a walk that starts and ends at the same site. However, the decomposition of a diagram into loops is ambiguous if there are self-intersections, that is, if there are vertices where four bonds meet. Let us consider, for example, the diagram in figure 5.15. It can be considered as the collection of two loops that meet at one vertex (a), or as a single loop whose path intersects itself (c) or does not (b). In order to obtain a nonambiguous sum, we assign to each diagram a factor $(-1)^n$, where n is the number of intersections. In this situation, the contribution of case (c) will be opposite to that of case (b), and they cancel out, leaving only the contribution of case (a). One can easily realize, then, that with this convention, the contribution of diagrams in which three bonds meet at a vertex identically vanishes, as can be seen in figure 5.16. In this way, the sum over all diagrams $\mathcal{G}$ is reduced to a sum over all loops, in which each loop appears with a weight proportional to $(-1)^n$, where n is the number of self-intersections. Notice that we do not allow vertices connected to only one bond and, therefore, the possibility that a walker retraces its steps backward.

We can now express the number of self-intersections of a loop by means of the following trick. It is well known that the total angle through which the tangent to the trajectory of a walker performing a loop turns around is given by $2\pi(\ell+1)$, where the parity of ℓ is equal to the parity of the number n of intersections. Thus, if we assign a factor $e^{i\phi/2}$ to each lattice point with turning angle ϕ, we have at the end of the loop a factor $(-1)^{\ell+1} = (-1)^{n+1}$, where n is the number of intersections. With this counting, each diagram made up of s loops will give a contribution proportional to $(-1)^{s+n}$. Thus we have to multiply this contribution by $(-1)^s$ in order to have the required sign in eq. (5.242).

In order to evaluate the angle ϕ, it is convenient to deal with directed loops. Let us denote by f_r the sum over all undirected loops consisting of r bonds (taking into account the factors t^r and $e^{i\phi n/2}$). Then, for example, the sum over all double loops of ℓ bonds is given by

$$\frac{1}{2}\sum_{r_1,r_2}\delta_{r_1+r_2,\ell} f_{r_1} f_{r_2},$$

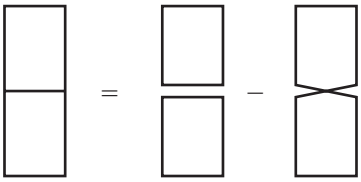


FIGURE 5.16. A diagram with a three-bond vertex can be obtained as the sum of two diagrams with numbers of intersections that differ by 1. Their contributions cancel out.

taking into account the possible permutations of the loops. Generalizing to all numbers of loops, and summing over all lengths ℓ, we obtain

$$\mathfrak{S} = \sum_{s=0}^{\infty} (-1)^s \frac{1}{s!} \sum_{r_1,\dots,r_s} f_{r_1} \cdots f_{r_s} = \exp\left(- \sum_{r=0}^{\infty} f_r \right). \tag{5.243}$$

In going from undirected to directed loops, each loop is encountered twice, and thus if we denote by v_r the sum of the contributions of directed loops with r bonds, we have

$$\mathfrak{S} = \exp\left(-\frac{1}{2} \sum_{r=0}^{\infty} v_r \right). \tag{5.244}$$

We now evaluate v_r. Given a lattice point (k, ℓ), let us denote the possible directions as follows:

$$\begin{aligned} &\mathrm{N}: \quad (k\ell) \longrightarrow (k, \ell+1), \\ &\mathrm{E}: \quad (k\ell) \longrightarrow (k+1, \ell), \\ &\mathrm{S}: \quad (k\ell) \longrightarrow (k, \ell-1), \\ &\mathrm{W}: \quad (k\ell) \longrightarrow (k-1, \ell). \end{aligned}$$

Let us denote by $W_r(k\ell\nu|k_0\ell_0\nu_0)$ the sum of all contributions of r-bond diagrams starting from lattice point (k_0, ℓ_0) in the direction $\nu_0 \in \{\mathrm{N}, \mathrm{E}, \mathrm{S}, \mathrm{W}\}$ and ending at lattice point (k, ℓ) in the direction ν. Each bond occurs with a factor $te^{i\phi/2}$, where ϕ is the change in direction in going to the next bond.

It is then possible to write a linear recursion relation for W_r:

$$W_{r+1}(k\ell\nu|k_0\ell_0\nu_0) = \sum_{k'\ell'\nu'} T_{k\ell\nu,k'\ell'\nu'} W_r(k'\ell'\nu'|k_0\ell_0\nu_0). \tag{5.245}$$

The transition matrix $\mathsf{T} = (T_{k\ell\nu,k'\ell'\nu'})$ has the expression

$$T_{k\ell\nu,k'\ell'\nu'} = tA_{\nu\nu'}\delta_{k',k+\alpha(\nu')}\delta_{\ell',\ell+\beta(\nu')}, \tag{5.246}$$

where

$$\begin{aligned} &\alpha(\mathrm{N}) = 0, && \beta(\mathrm{N}) = -1, \\ &\alpha(\mathrm{E}) = -1, && \beta(\mathrm{E}) = 0, \\ &\alpha(\mathrm{S}) = 0, && \beta(\mathrm{S}) = +1, \\ &\alpha(\mathrm{W}) = +1, && \beta(\mathrm{W}) = 0. \end{aligned}$$

The matrix $\mathsf{A} = (A_{\nu\nu'})$ (where $\nu, \nu' \in \{\mathrm{N}, \mathrm{E}, \mathrm{S}, \mathrm{W}\}$) is given by

$$\mathsf{A} = \begin{pmatrix} 1, & \omega, & 0, & \omega^* \\ \omega^*, & 1, & \omega, & 0 \\ 0, & \omega^*, & 1, & \omega \\ \omega, & 0, & \omega^*, & 1 \end{pmatrix}, \tag{5.247}$$

where

$$\omega = \mathrm{e}^{\mathrm{i}\pi/4}, \qquad \omega^* = \mathrm{e}^{-\mathrm{i}\pi/4}. \tag{5.248}$$

The connection between the weights W and the loop contributions v_r is given by

$$v_r = \frac{b_r}{r}, \tag{5.249}$$

where

$$b_r = \mathrm{Tr}\, W_r = \sum_{k\ell\nu} W_r(k\ell\nu|k\ell\nu). \tag{5.250}$$

The factor $1/r$ comes from the fact that a single diagram with r bonds can be obtained from r different walks, with different starting points. Now, from eq. (5.245), we have

$$\mathrm{Tr}\, W_r = \mathrm{Tr}\, \mathsf{T}^r = \sum_i \lambda_i^r, \tag{5.251}$$

where λ_i are the eigenvalues of the matrix T. From this equation, taking into account eqs. (5.244) and (5.249), we obtain

$$\mathfrak{S} = \exp\left(-\frac{1}{2}\sum_{r,i}\frac{1}{r}\lambda_i^r\right) = \exp\left[\frac{1}{2}\sum_i \ln(1-\lambda_i)\right] = \prod_i \sqrt{1-\lambda_i}. \tag{5.252}$$

Thus the problem boils down to the diagonalization of the matrix T. One can see from eq. (5.246) that T depends only on the differences in the indices k, ℓ. It can thus be diagonalized by a Fourier transformation. We set

$$T_{\nu\nu'}(m,n) = \sum_{k\ell} \mathrm{e}^{-2\pi\mathrm{i}(mk+n\ell)/L} T_{k\ell\nu,00\nu'}. \tag{5.253}$$

We then find

$$T_{\nu\nu'}(m,n) = t\begin{pmatrix} \gamma^*(n), & \omega\gamma^*(m), & 0, & \omega^*\gamma(m) \\ \omega^*\gamma^*(n), & \gamma^*(m), & \omega\gamma(n), & 0 \\ 0, & \omega^*\gamma^*(m), & \gamma(n), & \omega\gamma(m) \\ \omega\gamma^*(n), & 0, & \omega^*\gamma(n), & \gamma(m) \end{pmatrix}, \tag{5.254}$$

where

$$\gamma(m) = \mathrm{e}^{2\pi\mathrm{i}m/L}. \tag{5.255}$$

Thus, for given values of (m, n), we have

$$\begin{aligned}\prod_{i=1}^{4}[1-\lambda_i(m,n)] &= \det[\mathsf{I}-\mathsf{T}(m,n)] \\ &= (1+t^2)^2 - 2t(1-t^2)\left(\cos\frac{2\pi m}{L}+\cos\frac{2\pi n}{L}\right)^{1/2},\end{aligned} \tag{5.256}$$

where the product runs over L consecutive values of m and n.

Setting $p = 2\pi m/L$ and $q = 2\pi n/L$, the Helmholtz free energy is given by

$$\begin{aligned}F(T) = -Nk_{\mathrm{B}}T\Big\{&\ln 2 - \ln(1-t^2) \\ &+ \frac{1}{2}\int_{-\pi}^{+\pi}\frac{\mathrm{d}p\,\mathrm{d}q}{(2\pi)^2}\,\ln\left[(1+t^2)^2 - 2t(1-t^2)(\cos p+\cos q)\right]\Big\}.\end{aligned} \tag{5.257}$$

Let us consider the contribution of the integral. The minimum of the integrand is reached for $p = q = 0$ and is given by

$$\ln\left[(1+t^2)^2 - 4t(1-t^2)\right] = \ln\left(t^2+2t-1\right)^2. \tag{5.258}$$

The argument of the logarithm vanishes for

$$t = t_{\mathrm{c}} = \sqrt{2}-1, \tag{5.259}$$

which corresponds to the transition temperature T_{c} given by eq. (5.98):

$$\frac{J}{k_{\mathrm{B}}T_{\mathrm{c}}} = \frac{1}{2}\ln(1+\sqrt{2}). \tag{5.260}$$

In order to understand the behavior of F in the neighborhood of this temperature, let us introduce $\tau = t - t_{\mathrm{c}}$ and expand the integrand for small values of τ and p, q. One has

$$F(T) = \frac{1}{2}\int_{-\pi}^{+\pi}\frac{\mathrm{d}p\,\mathrm{d}q}{(2\pi)^2}\,\ln\left[c_1\tau^2 + c_2(p^2+q^2)\right] + \text{regular terms}, \tag{5.261}$$

where c_1 and c_2 are constants. Integrating, one obtains

$$F(T) = -a\tau^2\ln|\tau| + \text{regular terms}, \tag{5.262}$$

where $a > 0$ is a constant. The specific heat C is proportional to $-\mathrm{d}^2F/\mathrm{d}\tau^2$. We thus have

$$C \simeq a\ln|\tau| + \text{regular terms}, \tag{5.263}$$

indicating that the specific heat exhibits a logarithmic divergence at the critical temperature.

The evaluation of the spontaneous magnetization $m_0 = \langle\sigma_i\rangle$ proceeds in a similar way [Vdov65], starting from the relation

$$m_0^2 = \lim_{k\to\infty}\langle\sigma_{1,1}\sigma_{1,k}\rangle. \tag{5.264}$$

One obtains $m_0 = 0$ for $t < t_c$, given earlier, and

$$m_0 = \left[1 - \left(\frac{1/t - t}{2}\right)^4\right]^{1/8} \tag{5.265}$$

for $t > t_c$, that is, below the transition temperature. Thus, for small positive values of $\tau = t - t_c$, one has

$$m_0 \propto \tau^{\beta}, \tag{5.266}$$

where the exponent β is given by

$$\beta = \frac{1}{8}. \tag{5.267}$$

The connection between the Ising model and the statistics of loops can be interpreted more deeply as the equivalence between the model and a system of noninteracting fermions. This correspondence is exploited by Schulz, Mattis, and Lieb [Schu64] in their solution of the Ising model.

Recommended Reading

The phenomenology of phase transitions is explained in H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena*, Oxford, UK: Oxford University Press, 1971. The books devoted to phase transitions are now legion: it is sufficient to quote the series of volumes edited by C. Domb and M. S. Green (and later by C. Domb and J. L. Lebowitz), *Phase Transitions and Critical Phenomena*, London: Academic Press, 1971–2001. The mean-field and Landau theories are discussed in detail in L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed., part 1, Oxford, UK: Pergamon Press, 1980. A book with extensive discussion of phase transitions and order parameters is J. P. Sethna, *Statistical Mechanics: Entropy, Order Parameters, and Complexity*, 2nd ed., Oxford, UK: Oxford University Press, 2021. A simple presentation of the different ordered phases in condensed matter can be found in the excellent volume by P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics*, Cambridge, UK: Cambridge University Press, 1995. One book is entirely devoted to Landau phase transition theory in complex systems: J.-C.Tolédano and P. Tolédano, *The Landau Theory of Phase Transitions: Application to Structural, Incommensurate, Magnetic, and Liquid Crystal Systems*, Singapore: World Scientific, 1987.

6 Renormalization Group

Never do a calculation before you know the answer.
—J. A. Wheeler

The results of the preceding chapter imply that the statistics of fluctuations close to the critical point are very similar for the same system at different values of temperature and magnetic field (or, more generally, of those quantities that characterize distance from the critical point), provided one performs a rescaling of the lengths and of the order parameter. In 1966, Leo P. Kadanoff put forth an argument to explain this resemblance in the properties of fluctuations [Kada66]. Using an intuition by C. Di Castro and G. Jona-Lasinio [DiCa69] that the renormalization group (then a rather abstruse transformation in the theory of elementary particles) could have relevance for critical phenomena, K. G. Wilson deeply reformulated Kadanoff's idea in several papers of 1971 and in a review of 1972 [Wils72], establishing its connection with the particle theorists' renormalization group. Wilson's theory lies at the basis of the modern theory of critical phenomena. We are going to explain it, highlighting how the basic physical idea finds different implementations depending on the properties of the systems under study.

6.1 Block Transformation

Let us consider a two-dimensional Ising model in the proximity of the critical point. In the upper part of figure 6.1, you can see one of its typical configurations (the black squares represent an "up" spin). Since the distance from the critical point is small, the coherence length is large. We now define a new configuration using the majority rule: we group the spins together in 3×3 cells and assign an up spin to each cell if the majority of the spins that belong to it are up, and a down spin otherwise. We thus obtain a new configuration, which is represented in the lower portion of figure 6.1.

Kadanoff's idea is that this new configuration is a typical configuration of a two-dimensional Ising model, at a different (and, as we shall see, larger) distance from the critical point. To simplify, let us assume that the magnetic field vanishes. For reasons of symmetry, the magnetic field of the new configuration will also vanish. In this manner,

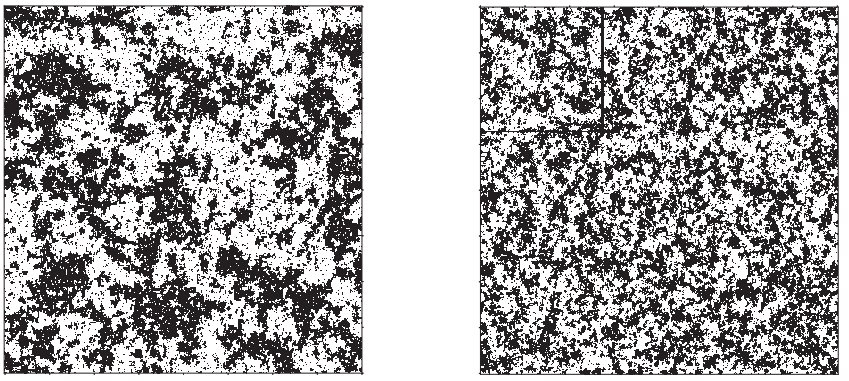


FIGURE 6.1. Block transformation. The left image represents a typical configuration of an Ising model on a square lattice at temperature $T = 1.05\, T_c$. From this configuration, one obtains the configuration in the right image by grouping the spins in the first configuration into 3×3 blocks and assigning to each block the spin value of the majority of the spins that compose it. The box at the top left in the right panel highlights the spins corresponding to those of the left panel. Inspired by [Wils79].

we will be dealing only with the t variable, proportional to $T - T_c$. We denote the value of this variable in the first configuration by t_0, and the corresponding value in the second by t_1. It is obvious that to each given value of t_0 corresponds a certain value of t_1. We have thus defined a transformation:

$$t_1 = \mathbb{T}(t_0). \tag{6.1}$$

By carrying out this transformation, the coherence length has changed. So if in the original model the spins were correlated up to a distance $\ell = \xi(t_0)$ (measured in units of the lattice constant), in the transformed model this distance will be equal to $\ell/3$, since each spin of the "new" model corresponds to 3^2 spins of the "old," and the new lattice step is therefore $b = 3$ times larger. Let us consider b as a variable, since we may in principle define other transformations with different values of b. We have therefore, in general, for a transformation grouping b^2 spins together,

$$\xi(t_1) = \frac{1}{b}\xi(t_0). \tag{6.2}$$

Since the new coherence length is smaller than the old, the new model is further away from the critical point than the old, as we had anticipated. The only exception occurs if we are exactly at the critical point, because in that case, the coherence length is infinite, and it must remain infinite also for the new model. The transformation $\mathbb{T}$ must therefore leave the critical point $t = 0$ unchanged, implying $\mathbb{T}(0) = 0$. We observe, moreover, that the transformation $\mathbb{T}$ is local, and that it requires considering the effects of a finite number of degrees of freedom each time. One can therefore expect it to be regular in its argument. Therefore, close to $t = 0$, we expect to have

$$\mathbb{T}(t) = \kappa\, t + \mathrm{O}\left(t^2\right), \tag{6.3}$$

where κ is a positive constant, larger than 1, since the transformed system is in general further away from the critical point.

We can repeat the same remarks by placing ourselves at the critical temperature and in a nonvanishing magnetic field, and therefore obtain the following expression for the transformation of h:

$$\mathbb{T}(h) = \kappa_h\, h + \mathrm{O}\left(h^3\right), \tag{6.4}$$

where $\kappa_h > 1$. For reasons of symmetry, only odd powers of h may appear. By combining everything, we obtain

$$\mathbb{T}(t, h) = (t', h'), \tag{6.5}$$

with $t' \simeq \kappa t$, $h' \simeq \kappa_h h$, while the corresponding transformation of the coherence length is given by

$$\xi(t', h') = \frac{1}{b}\xi(t, h). \tag{6.6}$$

These results imply the homogeneity of the thermodynamic functions and the scaling laws. To see it, let us place ourselves once more in a vanishing magnetic field, and let t tend to zero. We can now use the transformation to compare the value $\xi(t, h{=}0)$ of the coherence length with that of a system with a fixed distance t_0 from the critical point. We can iterate the transformation n times, so that $(t_0/t) = \kappa^n$. We then have

$$\xi\,(t, h{=}0) = \xi\,(\kappa^{-n} t_0, h{=}0) = b^n \xi(t_0, h{=}0). \tag{6.7}$$

This relation has the form of the generalized homogeneity relations that we discussed in chapter 5 if we set

$$\lambda = \kappa^{-n}, \qquad \lambda^{-\nu} = b^n. \tag{6.8}$$

The fact that λ was arbitrary in section 5.20, while our factors can assume only discrete values, is not essential. As a result of this comparison, we can express the critical exponent ν as a function of the constant κ:

$$\frac{1}{\nu} = \frac{\ln \kappa}{\ln b}. \tag{6.9}$$

Analogously, by considering the transformation at the critical temperature and in a vanishing field, we obtain

$$\frac{\beta\delta}{\nu} = \frac{\ln \kappa_h}{\ln b}. \tag{6.10}$$

These results were derived by Kadanoff.

Exercise 6.1 Derive the above relation.

In order to transform these observations into a theory that will allow for the effective calculation of the critical behavior, we must try to calculate the transformation $\mathbb{T}$. We then realize that the hypothesis that the new model is still an Ising model like the old one (although at a different distance from the critical point) is simplistic. If we want to describe the statistics of the fluctuations of the collective cell variables by means of a Hamiltonian, we are forced in general to introduce new interactions, such as interactions with several spins, or with next-to-nearest neighbors, and so on. It will therefore be necessary to consider the transformation $\mathbb{T}$ in a sufficiently large space of Hamiltonians. The key point is that in this transformation only short-range fluctuations are summed over, and therefore we expect that the transformation itself is non-singular. Thus it is, in principle, a regular mapping of the parameters that define the Hamiltonian, and we should be able to find effective approximation schemes. We devote the next sections to the definitions of some of these schemes.

6.2 Decimation in the One-Dimensional Ising Model

In this section, following the discussion in [Chan87, pp. 139ff], we provide an exact definition of the Kadanoff transformation $\mathbb{T}$ (or, rather, of one of its analogs) for the one-dimensional Ising model.

Consider a model defined by the Hamiltonian

$$\mathcal{H}(\sigma) = -\sum_i \left(J\sigma_i\sigma_{i+1} + h\sigma_i \right), \tag{6.11}$$

with $\sigma_i = \pm 1$, $\forall i$. It is useful to consider the expression of the partition function:

$$Z_N = \sum_\sigma \exp\left[\sum_i \left(K\sigma_i\sigma_{i+1} + \mu\sigma_i \right) \right], \tag{6.12}$$

where we have introduced the shorthands

$$K = \frac{J}{k_B T}, \qquad \mu = \frac{h}{k_B T}. \tag{6.13}$$

Let us now try to define a transformation in the spirit of the Kadanoff transformation $\mathbb{T}$ described in the preceding section. This transformation changes the fluctuation statistics of the system in an explicit way, dividing in particular the coherence length ξ by a known factor b, and providing an explicitly computable mapping $(K', \mu') = \mathbb{T}(K, \mu)$ of the parameters appearing in the model.

One transformation that allows us to obtain these results is **decimation**. We sum over the fluctuations of spins with an even index, and we obtain a model in which only spins with odd index appear—this is a new one-dimensional Ising model, with nearest-neighbor interactions, but in which the coherence length (measured in units equal to the lattice step) has been halved.

In order to perform this transformation explicitly, let us consider a spin with an even index i and its interaction with neighboring spins. For symmetry's sake, we assign half of

each spin's magnetic field term to its bond with the spin that precedes it, and half to that with the spin that follows it. We can then express the partition function as follows:

$$Z_N = \sum_{\{\sigma_{\text{odd}}\}} \prod_i \sum_{\sigma_{2i}} \exp\left[K(\sigma_{2i-1} + \sigma_{2i+1})\sigma_{2i} + \frac{\mu}{2}(\sigma_{2i-1} + 2\sigma_{2i} + \sigma_{2i+1})\right]. \tag{6.14}$$

By performing the sum over the even index spins, we easily obtain:

$$Z_N = \sum_{\{\sigma_{\text{odd}}\}} \exp\left\{\sum_i \left[\frac{\mu}{2}(\sigma_{2i-1} + \sigma_{2i+1}) + \ln 2\cosh\left[K(\sigma_{2i-1} + \sigma_{2i+1}) + \mu\right]\right]\right\}. \tag{6.15}$$

This is the partition function of a one-dimensional Ising model in disguise. In order to see this, let us first redefine the index:

$$2i - 1 \longrightarrow i, \qquad 2i + 1 \longrightarrow i + 1. \tag{6.16}$$

Let us also set

$$2\cosh\left[K(\sigma_i + \sigma_i) + \mu\right] = \zeta_0 \exp\left[K'\sigma_i\sigma_{i+1} + \frac{\delta\mu}{2}(\sigma_i + \sigma_{i+1})\right]. \tag{6.17}$$

The left-hand side can assume three values, depending on the values assumed by the two spins being considered; the right-hand side contains three parameters that have to be determined—in other words ζ_0, K', and $\delta\mu$. A solution can therefore exist. In fact, with a little algebra, one obtains:

$$\exp(4K') = \cosh(\mu + 2K)\cosh(\mu - 2K)\left(\cosh^2\mu\right)^{-1}; \tag{6.18a}$$

$$\exp(2\,\delta\mu) = \cosh(\mu + 2K)\left[\cosh(\mu - 2K)\right]^{-1}; \tag{6.18b}$$

$$\zeta_0 = 2\cosh\mu\ \exp(K'). \tag{6.18c}$$

We can now set $\mu' = \mu + \delta\mu$ and thus obtain the mapping $(K, \mu) \longrightarrow (K', \mu')$. As a result of this transformation, the partition function has been multiplied by $\zeta_0^{N/2}$. We can therefore write the effect of this transformation for $f = \ln Z_N/N$, which is proportional to the free energy per site:

$$f(K, \mu) = \frac{1}{2}\ln\zeta_0 + \frac{1}{2}f(K', \mu'). \tag{6.19}$$

On the other hand, the lattice step of the transformed model is twice that of the old, and the b factor is therefore equal to 2:

$$\xi(K', \mu') = \frac{1}{2}\xi(K, \mu). \tag{6.20}$$

In order to discuss the effects of the transformation, let us set $\mu = 0$. One then has $\mu' = 0$, as expected from symmetry considerations. The transformation of K takes the form

$$K' = \frac{1}{2}\ln\cosh 2K, \tag{6.21}$$

and one has

$$\zeta_0 = 2e^{K'} = 2\sqrt{\cosh 2K}. \tag{6.22}$$

It is easy to see that $K' \leq K$—one has equality only for $K = 0$. Therefore, starting from any finite value of K, the transformation takes us toward the fixed value $K = 0$, which corresponds to infinite temperature (and to vanishing coherence length). There is another value of K that remains unchanged under the transformation—namely, $K = \infty$ (which corresponds to vanishing temperature and infinite coherence length). The values of K that remain invariant define the **fixed points** of the transformation. Here the fixed point at $K = 0$ is stable, and that at $K = \infty$ is unstable, since the transformation tends to take the system away from it.

This behavior of the interactions under the transformation shows that there is no phase transition at a finite temperature. (Strictly speaking, one can talk of a transition at vanishing temperature.) We now show that we can use this transformation in order to evaluate the free energy.

In order to have an effective calculation scheme, it is useful to express the free energy calculated in K as a function of that calculated in K', by making use of the equation

$$f(K) = \frac{1}{2} f(K') + \frac{1}{2} \ln 2 + \frac{K'}{2}, \tag{6.23}$$

which derives from the expressions of ζ_0 and K' obtained earlier. We then have to invert the relation between K and K', so as to have

$$K = \mathbb{T}^{-1}(K') = \frac{1}{2} \cosh^{-1}\left(e^{2K'}\right). \tag{6.24}$$

In this manner, possible errors made when estimating $f(K')$ will be compressed, since they are multiplied by a factor of $1/2$ at each iteration.

Our strategy is to start from a specific value K_0 of K, and from a corresponding initial estimate of $f(K_0)$. By making use of eq. (6.24) we obtain a sequence $K_1, K_2, \ldots$, where $K_n = \mathbb{T}^{-1}(K_{n-1})$. The relation (6.23) then gives us the corresponding value of $f(K_n)$. Since K grows at each iteration, we can start with a very small value—for example, 0.01—where we can estimate $f(0.01) \simeq \ln 2$, which is the value for $K = 0$. By iterating, we obtain table 6.1, in which the first column contains the successive values of K_n, the second column contains the corresponding values of $f(K_n)$ estimated by means of the transformation, and the third column contains the values calculated from the exact solution. We can see that a moderate computational investment allows us to obtain almost exact free energy values.

6.3 Two-Dimensional Ising Model

Let us now consider a two-dimensional Ising model in a vanishing magnetic field, defined over a square lattice. As with the one-dimensional model, we shall use the notation $K = J/k_B T$.

Table 6.1 Decimation results for the free energy in the one-dimensional Ising model.

K	$f(K)$ (6.23)	$f(K)$ (exact)
0.01	0.693147	0.693197
0.100334	0.698147	0.698172
0.327447	0.745814	0.745827
0.636247	0.883204	0.883210
0.972710	1.106299	1.106302
1.316710	1.386078	1.386080
1.662637	1.697968	1.697968
2.009049	2.026876	2.026877
2.355582	2.364536	2.364536
2.702146	2.706633	2.706633

Note: In the first column, $K = J/k_B T$ for successive iterations of eq. (6.24), and in the second column, $f(K)$ is the corresponding value of free energy per spin, calculated by means of eq. (6.23), and compared with the exact value $f(K) = \ln 2 \cosh K$. After [Chan87, p. 142].

The spins are arranged on two intertwined sublattices, such that each site that belongs to one sublattice is surrounded by spins belonging to the other sublattice. The sublattices are distinguished by the sum of the indices corresponding to the two coordinates—if this sum is even, the site belongs to the "even" sublattice; otherwise, it belongs to the "odd" one. The new model is also defined over a square lattice (rotated by 45 degrees), while the lattice step is increased by a factor $b = \sqrt{2}$.

Let us focus our attention on one element (figure 6.2), composed of a spin (which we will denote by σ_0) and its four neighbors. The contribution of this element to the partition function is given by

$$\sum_{\sigma_0} \exp\left[K\sigma_0\left(\sigma_1+\sigma_2+\sigma_3+\sigma_4\right)\right] = 2\cosh\left[K\left(\sigma_1+\sigma_2+\sigma_3+\sigma_4\right)\right]. \tag{6.25}$$

We would like to write this expression in the form

$$\exp\left[\frac{K'}{2}\left(\sigma_1\sigma_2+\sigma_2\sigma_3+\sigma_3\sigma_4+\sigma_4\sigma_1\right)\right],$$

maybe up to a constant factor. The factor $1/2$ is due to the fact that the 1-2, 2-3, 3-4, 4-1 bonds appear twice in the new model—once in the cell we are considering, and a second time in each of the neighboring cells. One cannot, however, obtain this expression by means of a single interaction constant. Indeed, the right-hand side of eq. (6.25) is symmetrical with respect to all the permutations of (1,2,3,4), while the interaction among nearest neighbors is not. On the other hand, the result of the decimation assumes only three different values, and it should be possible therefore to parametrize it with two coupling constants and one constant factor. It turns out that we need three coupling constants and one factor.

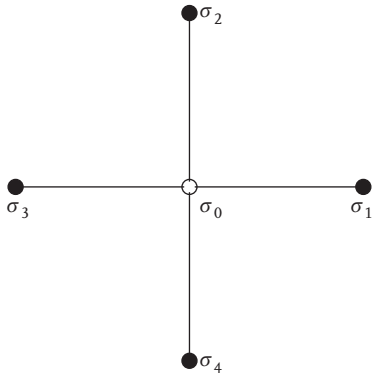


FIGURE 6.2. Decimation in the $2d$ Ising model. The "decimating" spin σ_0 and its four neighbors.

It is therefore necessary to introduce other interactions into the new model—in our case, we will need the next-to-nearest-neighbor interactions (σ_1-σ_3 and σ_2-σ_4, for example), represented by a parameter K_1, and a four-spin interaction, represented by a parameter L. We therefore set

$$\begin{aligned} 2\cosh\left[K\left(\sigma_1+\sigma_2+\sigma_3+\sigma_4\right)\right] &= \zeta_0(K) \\ &\times \exp\left[\frac{K'}{2}\left(\sigma_1\sigma_2+\sigma_2\sigma_3+\sigma_3\sigma_4+\sigma_4\sigma_1\right)+\right. \\ &\left.+K_1\left(\sigma_1\sigma_3+\sigma_2\sigma_4\right)+L\,\sigma_1\sigma_2\sigma_3\sigma_4\right]. \end{aligned} \tag{6.26}$$

One thus obtains the equations

$$2\cosh 4k = \zeta_0(K)\exp\left(2K'+2K_1+L\right); \tag{6.27a}$$

$$2\cosh 2K = \zeta_0(K)\exp\left(-L\right); \tag{6.27b}$$

$$2 = \zeta_0(K)\exp\left(-2K_1+L\right); \tag{6.27c}$$

$$2 = \zeta_0(K)\exp\left(-2K'+2K_1+L\right). \tag{6.27d}$$

It is a simple algebra exercise to show that

$$K' = \frac{1}{4}\ln\cosh 4K; \tag{6.28a}$$

$$K_1 = \frac{1}{8}\ln\cosh 4K; \tag{6.28b}$$

$$L = \frac{1}{8}\ln\cosh 4K - \frac{1}{2}\ln\cosh 2K; \tag{6.28c}$$

$$\zeta_0(K) = 2\left(\cosh 2K\right)^{1/2}\left(\cosh 4K\right)^{1/8}. \tag{6.28d}$$

Even if we start from a model in which only nearest-neighbor interactions are present, as a result of the first decimation we also obtain next-nearest-neighbor and four-spin interactions. A second decimation would introduce even more interactions. In order to be

able to implement our program, we need to simplify our result so that it can be represented by a smaller number of interaction parameters.

One possibility is to disregard all interactions but the nearest-neighbor ones. The resulting equation is

$$K' = \frac{1}{4} \ln \cosh 4K, \tag{6.29}$$

which looks essentially like the transformation in the one-dimensional Ising model and, as in that case, does not describe a phase transition.

In order to improve our approximation, we must somehow also take the other interactions into account. We are not able to iterate the transformation if other interactions appear in our Hamiltonian, such as the next-nearest-neighbor ones, for instance. We must therefore find a way to project the transformation (6.28) over the nearest-neighbor interactions axis, but take the effect of the other interactions at least partially into account. There is no systematic way of proceeding, and the procedure will thus be validated by the results we find. One choice that has been found acceptable is to neglect the four-body interaction and to choose the nearest-neighbor interaction coefficient so that the energy of the ground state is equal to that due to K' and K_1:

$$K' \sum_{<ij>} \sigma_i \sigma_j + K_1 \sum_{(\ell m)} \sigma_\ell \sigma_m \simeq \overline{K'} \sum_{<ij>} \sigma_i \sigma_j. \tag{6.30}$$

The first sum runs over the pairs of nearest neighbors, and so does that on the right-hand side; the second sum on the left-hand side runs over the pairs of next-nearest neighbors.

By comparing the two sides of this equation in the state in which all spins are parallel, we obtain

$$NK' + NK_1 \simeq N\overline{K'}, \tag{6.31}$$

and, therefore,

$$\overline{K'} \simeq K' + K_1 = \frac{3}{8} \ln \cosh 4K. \tag{6.32}$$

The approximate transformation has the expression (with an obvious change in notation)

$$K' = \frac{3}{8} \ln \cosh 4K; \tag{6.33a}$$

$$f(K') = 2f(K) - \ln \left[2 \left(\cosh 2K \right)^{1/2} \left(\cosh 4K \right)^{1/8} \right]. \tag{6.33b}$$

In figure 6.3, we plot K' as a function of K, as obtained from eq. (6.33a). This relation admits one fixed point, which solves the equation

$$K^* = \frac{3}{8} \ln \cosh 4K^*. \tag{6.34}$$

The solution is

$$K^* \simeq 0.506981. \tag{6.35}$$

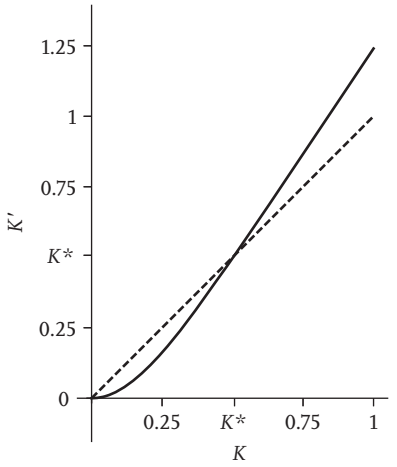


FIGURE 6.3. Renormalized interaction parameter K' as a function of K for the two-dimensional Ising model, according to eq. (6.33a). The dashed line corresponds to $K' = K$. Notice the fixed point at $K^* \simeq 0.506981$.

As we see from fig. 6.3, for $K < K^*$ we have $K' < K$, whereas for $K > K^*$ we have $K' > K$. Therefore, if $K \neq K^*$, the iteration pulls K away from the fixed point, which is therefore *unstable*, and we can identify K^* with the critical value K_c of $J/k_B T$. This compares reasonably well with the exact result $J/k_B T_c \simeq 0.440687$ given by eq. (5.98). For $K = K^* + \delta K$, we have

$$K' = K^* + \kappa\, \delta K + \mathrm{O}\left(\delta K^2\right), \tag{6.36}$$

where

$$\kappa = \left.\frac{\mathrm{d}K'}{\mathrm{d}K}\right|_{K^*} \simeq 1.44892. \tag{6.37}$$

This quantity is greater than 1, and thus the iterates move away from the fixed point, as we just argued from fig. 6.3. We are therefore in the situation described by Kadanoff's argument.

Exercise 6.2 Solve numerically eq. (6.34). (Hint: How do you transform a fixed point from unstable into stable?)

Since if $K = K^*$ we remain at the fixed point, we can assume that δK is proportional to the distance from the transition point, and that it thus plays the role of t in Kadanoff's argument. In this fashion, we have realized, if only approximately, Kadanoff's program for the two-dimensional Ising model, and we have obtained an estimate of the exponent ν,

$$\frac{1}{\nu} = \frac{1}{\ln b} \ln \left.\frac{\mathrm{d}K'}{\mathrm{d}K}\right|_{K^*} \simeq 1.06996, \tag{6.38}$$

which is an excellent approximation of the exact result $1/\nu = 1$, especially if compared with the classical result $1/\nu = 2$. We have expressed the results as a function of $1/\nu$, because as one can see from our equations, this is the quantity that is obtained most directly from the calculation.

6.4 Relevant and Irrelevant Operators

As we have seen, the transformation from the original model to the model with collective variables introduces in general new interactions. At least in principle, we must therefore define the Kadanoff transformation $\mathbb{T}$ in a space that contains all the possible interactions.

Let us suppose that we deal with a model defined over a cubic lattice in d dimensions, and in which a real variable ϕ_i is associated with each lattice point i. In general, the partition function has the expression

$$Z = \int \prod_i \mathrm{d}\phi_i \, \exp\left[-\mathcal{H}(\phi)\right], \tag{6.39}$$

where $\phi = (\phi_i)$ and

$$\mathcal{H}(\phi) = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{i_1,\ldots,i_n} K^{(n)}_{i_1\ldots i_n} \phi_{i_1} \cdots \phi_{i_n}. \tag{6.40}$$

It is clear from these equations that the temperature T has been absorbed in the definition of the interaction coefficients $K = \left(K^{(n)}_{i_1\ldots i_n}\right)$. We consider in the following, except when explicitly stated otherwise, only Hamiltonians that are invariant with respect to the transformation $\phi \longrightarrow -\phi$. Therefore, only even values of n appear in this expansion. Since the transformation $\mathbb{T}$ conserves this symmetry, this is also true for the transformed Hamiltonians.

Thus the model is identified by the collection $K = \left(K^{(n)}_{i_1\ldots i_n}\right)$ of the interaction coefficients. We can consider them as coordinates in an infinite-dimensional space of all possible Hamiltonians of the form given by eq. (6.40). We can then define the block transformation $\mathbb{T}$ as follows:

1. Identify nonoverlapping cells ℓ of neighboring lattice points. Each cell contains b^d points. For each cell ℓ, define the block variable Φ_ℓ by $\Phi_\ell = \sum_{i\in\ell} \phi_i$.

2. Express the effective interaction in the transformed model by summing over the original variables $\phi = (\phi_i)$, constrained by the given values of the block variables $\Phi = (\Phi_\ell)$. This corresponds to expressing the effective Hamiltonian $\overline{\mathcal{H}}(\Phi)$ by

$$\zeta_0^{N/b^d} \exp\left[-\overline{\mathcal{H}}(\Phi)\right] = \int \prod_i \mathrm{d}\phi_i \prod_\ell \delta\left(\Phi_\ell - \sum_{i\in\ell} \phi_i\right) \exp\left[-\mathcal{H}(\phi)\right], \tag{6.41}$$

   where the factor ζ_0^{N/b^d} takes into account the constant factors arising from the integration in each cell, allowing us to keep the relation $\overline{\mathcal{H}}(0) = 0$.

3. Redefine the lattice constant, so as to have a variable Φ_ℓ for each point ℓ of the cell lattice—the distances will be therefore multiplied by a factor of $1/b$.

4. Define the new variable $\phi'_\ell = \zeta \Phi_\ell$ (by suitably rescaling the collective variable Φ_ℓ). The factor ζ is chosen so that the transformation has a nontrivial fixed point. Practically speaking, it is sufficient to require that the interactions one obtains decrease fairly rapidly with the distance.

In this fashion, we have defined a transformation $\mathbb{T}$ between the Hamiltonians:

$$\mathcal{H}'(\phi') = (\mathbb{T}\mathcal{H})(\phi). \tag{6.42}$$

By defining $f(\mathcal{H}) = \ln Z(\mathcal{H})/N$, the properties of the transformation can be summarized by the relations

$$f(\mathcal{H}) = \left(N/b^d\right)\left[\ln \zeta_0(\mathcal{H}') + f(\mathcal{H}')\right]; \tag{6.43a}$$

$$\xi(\mathcal{H}) = b \cdot \xi(\mathcal{H}'); \tag{6.43b}$$

$$\langle \phi_0 \phi_{\mathbf{r}} \rangle_{\mathcal{H}} = b^{2d} \zeta^{-2} \left\langle \phi'_0 \phi'_{\mathbf{r}/b} \right\rangle_{\mathcal{H}'}. \tag{6.43c}$$

In the third relation, we are evaluating the correlation between variables situated at the origin and at the position $\boldsymbol{r}$ in space, and we have taken into account the fact that the new variables ϕ' are proportional to the sum of b^d old variables.

In the space of the Hamiltonians, we can identify the **critical surface** S_c defined by $\xi(\mathcal{H}) = \infty$, $\forall \mathcal{H} \in S_c$. This surface is invariant with respect to $\mathbb{T}$, as one can see from eq. (6.43b). Starting from any point on this surface and iterating the transformation, we expect to end up at one of the fixed points $\mathcal{H}^*$ of the transformation, which satisfies the equation

$$\mathbb{T}\mathcal{H}^* = \mathcal{H}^*. \tag{6.44}$$

Keeping eq. (6.43c) in mind, we see that for each Hamiltonian belonging to the critical surface, the correlation functions relative to large distances are eventually proportional to those characteristic of $\mathcal{H}^*$. This is the microscopic justification of universality: the properties of the critical fluctuations are determined by $\mathcal{H}^*$—in other words, by the transformation $\mathbb{T}$—rather than by the details of the system's Hamiltonian (which only plays the role of the initial condition for the iteration procedure).

The critical surface S_c has codimension 1 because it is defined by a single equation. For our initial model (for example, the Ising model), the Hamiltonian, as the temperature varies, runs over a line. If we also introduce an applied magnetic field h, the Hamiltonian runs over a two-dimensional surface. We call the surface S_0 spanned by the Hamiltonian of our model the **elementary surface**. Generically, the elementary surface meets the critical surface in a point $\mathcal{H}_c$, which represents the critical Hamiltonian. This situation is schematically represented in figure 6.4.

Starting from $\mathcal{H}_c$ and iterating the transformation, we reach the fixed point $\mathcal{H}^*$. Iterating the transformation at this point does not change the Hamiltonian. We can now express the effect of the transformation on the correlation functions:

$$\langle \phi_0 \phi_{\mathbf{r}} \rangle_{\mathcal{H}^*} = b^{-2d} \zeta^{-2} \left\langle \phi_0 \phi_{\mathbf{r}/b} \right\rangle_{\mathcal{H}^*}. \tag{6.45}$$

By transforming according to Fourier, we obtain

$$C(\boldsymbol{k}, \mathcal{H}^*) = b^{-d} \zeta^{-2} C(\boldsymbol{k}b, \mathcal{H}^*). \tag{6.46}$$

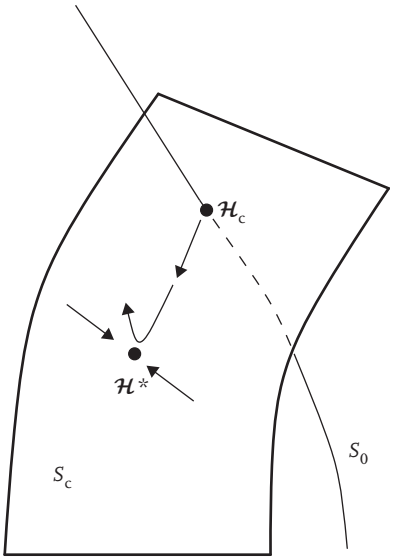


FIGURE 6.4. Scheme of the critical surface S_c and of the elementary surface S_0 in the space of the Hamiltonians. The critical Hamiltonian $\mathcal{H}_c$ is at the intersection of S_0 with S_c. The fixed point $\mathcal{H}^*$ of the Kadanoff transformation $\mathbb{T}$ is also shown. The arrows point in the direction of the flux under the iteration of the Kadanoff transformation. If one starts from a point on S_0 slightly above S_c, the flow first approaches $\mathcal{H}^*$, then goes away from it in a well-defined direction, as represented by the bent arrow.

This is a generalized homogeneity relation, from which we obtain

$$d + 2 - \eta = -2\frac{\ln \zeta}{\ln b}. \tag{6.47}$$

This result also shows us that the ζ factor is necessary. Indeed, setting $\zeta = 1$ implies $\eta = d + 2$, which does not hold in general.

Let us now consider our model slightly above the critical temperature, by setting

$$\mathcal{H}_0 = \mathcal{H}_c + \delta\mathcal{H}, \tag{6.48}$$

where $\delta\mathcal{H}$ is proportional to the distance from the critical temperature. By continuity, we expect that the iteration of the transformation $\mathbb{T}$ will at first bring us into proximity to the fixed point $\mathcal{H}^*$, so that, for a suitable number n_0 of iterations, we have

$$\mathbb{T}^{n_0}\mathcal{H}_0 = \mathcal{H}^* + \delta\mathcal{H}^*, \tag{6.49}$$

where $\delta\mathcal{H}^*$ is small. By applying the transformation once more, we obtain

$$\mathbb{T}\,(\mathcal{H}^* + \delta\mathcal{H}^*) \simeq \mathcal{H}^* + \mathcal{L}\,\delta\mathcal{H}^*, \tag{6.50}$$

where $\mathcal{L}$ is a linear operator. We can express $\delta\mathcal{H}^*$ in terms of the right eigenvectors of $\mathcal{L}$:

$$\delta\mathcal{H}^* = \sum_\ell c_\ell\,\delta\mathcal{H}^{(\ell)}, \tag{6.51}$$

where

$$\mathcal{L}\,\delta\mathcal{H}^{(\ell)} = \lambda^{(\ell)}\,\delta\mathcal{H}^{(\ell)}. \tag{6.52}$$

It follows from our hypotheses that $\mathcal{L}$ admits only one eigenvector greater than 1. Before "proving" this, let us suppose that it is true and examine its consequences. If $\delta\mathcal{H}^*$ is given

by eq. (6.51), we have

$$\mathcal{L}\,\delta\mathcal{H}^* = \sum_\ell \lambda^{(\ell)} c_\ell\,\delta\mathcal{H}^{(\ell)}. \tag{6.53}$$

If $\lambda^{(1)} > 1$, and $\left|\lambda^{(\ell)}\right| < 1$ for $\ell > 1$, the first term in this sum is larger than that in eq. (6.51), while each of the following terms is smaller. Now, if we start from a Hamiltonian in S_c and iterate $\mathbb{T}$ indefinitely, we eventually reach $\mathcal{H}^*$. This implies that for $\mathcal{H} \in S_c$ the coefficient c_1 of $\delta\mathcal{H}^{(1)}$ in eq. (6.51) must vanish. Since the mapping from $\mathcal{H}$ to $\mathbb{T}^{n_0}\mathcal{H}$ is regular, this implies that c_1, which vanishes at the critical temperature T_c, is proportional to the distance from T_c:

$$c_1 \propto (T - T_c)\,. \tag{6.54}$$

While successive iterations make c_1 increase, they also make all the other terms in the sum decrease. Thus the terms other than the first in eq. (6.51) do not affect the behavior near the critical point, and are therefore called **irrelevant** perturbations. Other models whose fundamental surface crosses the same critical surface are governed by the same fixed point (and, as we shall see, have the same critical exponents).

As a first approximation, we can disregard the irrelevant perturbations and consider the behavior of $\mathcal{H}(c_1) = \mathcal{H}^* + c_1\,\delta\mathcal{H}^{(1)}$. We thus obtain

$$\xi\left(\mathcal{H}(c_1)\right) = b\,\xi\left(\mathcal{H}(\lambda^{(1)} c_1)\right). \tag{6.55}$$

Recalling eq. (6.54), we can compare this result with the expected behavior $\xi \propto |T - T_c|^{-\nu}$, and we obtain the expression of the exponent $1/\nu$:

$$\frac{1}{\nu} = \frac{\ln \lambda^{(1)}}{\ln b}. \tag{6.56}$$

The other eigenvalues of $\mathcal{L}$ govern the approach to the fixed point, and therefore the corrections to the critical behavior.

The set of all Hamiltonians $\mathcal{H}$ such that $\lim_{n\to\infty} \mathbb{T}^n\mathcal{H} = \mathcal{H}^*$ is called the **attraction basin** of $\mathcal{H}^*$. In the simple case we considered, this coincides with the critical surface S_c. Let us now consider a generic Hamiltonian $\mathcal{H}$ close to a fixed point $\mathcal{H}^*$ (not necessarily the one we discussed above). We can expand $\mathcal{H}$ as in eq. (6.51). Upon applying the transformation $\mathbb{T}$, some terms will grow: they are called the **relevant** perturbations (or operators), because they take the iterates away from the fixed point $\mathcal{H}^*$. When these terms are present, the asymptotic critical behavior is no more described by $\mathcal{H}^*$. In our case, there is only one relevant operator, namely $\delta\mathcal{H}^{(1)}$. However, adding to the Hamiltonian a perturbation that explicitly breaks the symmetry $\phi \longrightarrow -\phi$ introduces at least one additional relevant operator. As for the terms that grow smaller as the transformation $\mathbb{T}$ is applied, they do not modify the asymptotic critical behavior, which justifies the term "irrelevant perturbations." We can also see that the codimension of the basin of attraction of $\mathcal{H}^*$ is equal to the number of relevant operators in the expansion appearing in eq. (6.51). Thus, if this codimension

is equal to 1, as we have assumed, there can be one, and only one, relevant operator, as we had anticipated.

On the other hand, it is possible that the critical surface S_c contain several attraction basins, relative to different fixed points of the transformation $\mathbb{T}$. This often takes place when the Hamiltonian has an additional symmetry that can be explicitly broken. In this case there will be one (or more) fixed points with the broken symmetry, with an attraction basin of codimension 1, and a fixed point with the full symmetry, with an attraction basin of codimension 2, since the symmetry-breaking operator is relevant close to the symmetric fixed point. We discuss a case of this kind in sec. 6.7.

6.5 Finite Lattice Method

In this section, we discuss a technique that can be used (in principle) to obtain a sequence of approximations to the exact Kadanoff transformation $\mathbb{T}$. For simplicity, and following the authors of the method [Niem76], we consider a two-dimensional Ising model over a triangular lattice. Universality guarantees that the critical exponents of this model are the same as those of the usual Ising model over a square lattice.

We want define the Kadanoff transformation in the following way: one arranges the spins from the original model into groups of three, according to the sketch shown in figure 6.5. One then associates a block-spin μ_ℓ with each block ℓ. These spins are also arranged over a triangular lattice, with a lattice step equal to $\sqrt{3}$ times the original constant.

Given the Ising Hamiltonian of the original model,

$$\mathcal{H}(\sigma; K, h) = -K \sum_{<ij>} \sigma_i \sigma_j - h \sum_i \sigma_i, \tag{6.57}$$

we obtain the new Hamiltonian $\mathcal{H}'$ by summing over the spins σ_i according to the following rule:

$$e^{-Nf_0(K,h) - \mathcal{H}'(\mu;K',h')} = \sum_\sigma \mathcal{P}(\sigma, \mu)\, e^{-\mathcal{H}(\sigma;K,h)}. \tag{6.58}$$

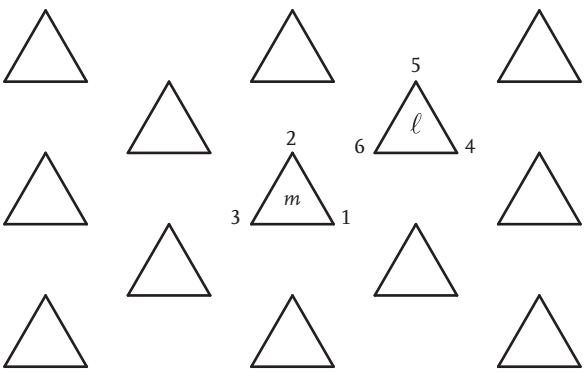


Figure 6.5. Block-spin transformation in a triangular lattice. The numbers identify the spins of the original model and the letters identify the blocks.

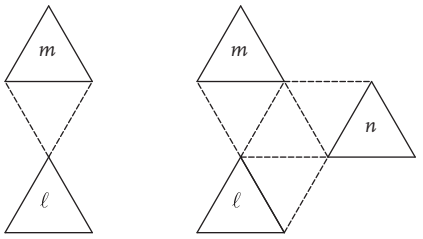


FIGURE 6.6. Finite size approximation. Left: A finite lattice with two blocks and six spins. Right: A finite lattice with three blocks and nine spins.

In this equation, the **projector** $\mathcal{P}$ represents the local relation that holds between the σs and the μs and is chosen in order to satisfy the normalization

$$\sum_{\mu} \mathcal{P}(\sigma, \mu) = 1, \qquad \forall \sigma, \tag{6.59}$$

so as to guarantee the invariance of the partition function under the transformation (6.58).

Choosing the projector $\mathcal{P}$ is equivalent to choosing the relation between the σs and the μs. One possible choice is the *majority rule*:

$$\mu_\ell = \phi_\ell(\sigma) = \operatorname{sign} \sum_{i \in \ell} \sigma_i. \tag{6.60}$$

In this manner, $\mathcal{P}$ admits the expression

$$\mathcal{P}(\sigma, \mu) = \prod_{\ell} \delta_{\mu_\ell, \phi_\ell(\sigma)}. \tag{6.61}$$

Although expression (6.58), given (6.61), is explicit, this does not imply that it is actually computable. We discuss just one of many possible methods, in which one considers explicitly a finite lattice, for which the sum (6.58) can be calculated exactly. Let us therefore consider two blocks of three spins, ℓ and m, nearest neighbors of each other, as shown in figure 6.6, left.

To simplify, let us assume that we are in a vanishing magnetic field. The recurrence relation takes the form

$$\mathrm{e}^{-f_0(K)+K'\mu_\ell\mu_m} = \sum_{\sigma} \mathcal{P}(\sigma, \mu)\, \mathrm{e}^{-\mathcal{H}(\sigma)}. \tag{6.62}$$

By making the dependence on μ explicit, we obtain

$$\mathrm{e}^{-f_0(K)+K'} = \sum_{\sigma} \delta_{1,\phi_\ell(\sigma)} \delta_{1,\phi_m(\sigma)}\, \mathrm{e}^{-\mathcal{H}(\sigma)}; \tag{6.63a}$$

$$\mathrm{e}^{-f_0(K)-K'} = \sum_{\sigma} \delta_{1,\phi_\ell(\sigma)} \delta_{-1,\phi_m(\sigma)}\, \mathrm{e}^{-\mathcal{H}(\sigma)}. \tag{6.63b}$$

We can calculate these expressions by brute force, summing over all the configurations of spins allowed by the delta function. One thus obtains

$$\mathrm{e}^{-f_0(K)+K'} = \mathrm{e}^{8K} + 3\mathrm{e}^{4K} + 2\mathrm{e}^{2K} + 3 + 6\mathrm{e}^{-2K} + \mathrm{e}^{-4K}; \tag{6.64a}$$

$$\mathrm{e}^{-f_0(K)-K'} = 2\mathrm{e}^{4K} + 2\mathrm{e}^{2K} + 4 + 6\mathrm{e}^{-2K} + 2\mathrm{e}^{-4K}. \tag{6.64b}$$

This implies

$$K' = \frac{1}{2} \ln \frac{e^{8K} + 3e^{4K} + 2e^{2K} + 3 + 6e^{-2K} + e^{-4K}}{2e^{4K} + 2e^{2K} + 4 + 6e^{-2K} + 2e^{-4K}}. \tag{6.65}$$

This recurrence relation admits one fixed point $K^* = 0.365297$, with $dK'dK|_{K^*} = 1.54348$, which yields the exponent $1/\nu = 0.790165$. Although these results are not very good, one can systematically improve the approximation by considering larger lattices [Niem76]. For example, considering the three-block lattice shown in fig. 6.6, right, one obtains the recursion relation

$$\begin{aligned} K' = \frac{1}{4} \ln \big[& \big(6\,e^{-5K} + 11\,e^{-3K} + 12\ e^{-K} + 9\,e^{K} + 7\,e^{3K} + 9\,e^{5K} \\ & + 3\,e^{7K} + 3\,e^{9K} + 3\,e^{11K} + e^{15K}\big) \\ & \Big/ \big(2\,e^{-7K} + 7\,e^{-5K} + 15\,e^{-3K} + 15\,e^{-K} + 9\,e^{K} + 9\,e^{3K} \\ & + 5\,e^{5K} + e^{7K} + e^{9K}\big)\big]. \end{aligned} \tag{6.66}$$

This relation exhibits a fixed point at $K^* = 0.304689$, with $dK'/dK|_{K^*} = 1.49411$, corresponding to an exponent $1/\nu = 0.730974$. (The results of the original paper are slightly different.) On the other hand, if one considers the effects of a magnetic field on the same finite lattice, one needs to introduce both a one-spin and a three-spin interaction. When dealing with still larger lattices, all the interactions compatible with the symmetry of the system will come into play one after another. It is therefore necessary to broaden the space of the interactions being considered.

6.6 Renormalization in Fourier Space

In the following sections, we will describe how one can obtain the universal critical properties by means of a version of the Kadanoff transformation in Fourier space. These developments closely follow Wilson's original formulation, which is discussed in [Wils72]. It may also be useful to read the books by D. Amit [Amit05] and N. Goldenfeld [Gold92].

Here, the point of view is quite different from that in the previous sections. Instead of considering specific models, we take advantage of universality in order to define some Hamiltonians that possess only the essential properties that determine the critical behavior. This approach leads to approximate expressions of the critical exponents, in the form of an expansion in which only these essential properties appear: the dimensionality of space, d, and that of the order parameter, n, plus the symmetries obeyed by the system.

Let us consider an n-component generalization of the Heisenberg model, defined by the Hamiltonian

$$\mathcal{H}(\phi) = -\frac{1}{2} \sum_{ij} K_{ij} \boldsymbol{\phi}_i \cdot \boldsymbol{\phi}_j - \sum_i \boldsymbol{h} \cdot \boldsymbol{\phi}_i, \tag{6.67}$$

where $\phi = (\boldsymbol{\phi}_i)$ and, for each lattice point i, we have

$$\boldsymbol{\phi}_i = \left(\phi_i^\alpha\right), \qquad \phi_i^\alpha \in \mathbb{R}, \qquad \sum_{\alpha=1}^{n} \left(\phi_i^\alpha\right)^2 = |\boldsymbol{\phi}|^2 = 1. \tag{6.68}$$

(Of course, for $n = 1$, we recover a model of the Ising type.) For convenience, we have included the factor $1/k_B T$ in the interaction parameters K and $\boldsymbol{h}$, and the Boltzmann factor is therefore simply given by $e^{-\mathcal{H}(\phi)}/Z$, where the partition function Z is given by

$$Z = \mathrm{Tr}\, e^{-\mathcal{H}(\phi)}. \tag{6.69}$$

Here the trace Tr is defined by

$$\mathrm{Tr}\, \mathcal{A}(\phi) = \int \prod_i \left[\prod_{\alpha=1}^{n} d\phi_i^\alpha \, \delta\left(\sum_{\alpha=1}^{n} \left(\phi_i^\alpha\right)^2 - 1 \right) \right] \mathcal{A}(\phi), \tag{6.70}$$

where $\mathcal{A}(\phi)$ is a generic function of the configuration $\phi = (\boldsymbol{\phi}_i)$. It is often useful to slightly relax the constraint over the spin magnitude and to consider a more general formulation in which

$$\mathrm{Tr} = \int \prod_i \left[\prod_{\alpha=1}^{n} d\phi_i^\alpha \, W(\boldsymbol{\phi}_i) \right], \tag{6.71}$$

where $W(\boldsymbol{\phi})$ is a given function, called the **weight**. While in eq. (6.70) we have $W(\boldsymbol{\phi}) = \delta\left(|\boldsymbol{\phi}|^2 - 1\right)$, another possible choice for W is the following:

$$W(\boldsymbol{\phi}) = \mathcal{N}\, e^{-u\left(|\boldsymbol{\phi}|^2 - 1\right)^2}, \tag{6.72}$$

where $\mathcal{N}$ is a normalization constant. This function has a maximum for $|\boldsymbol{\phi}| = 1$ and decreases more and more rapidly the larger u becomes. Expression (6.70) therefore can be recovered in the limit $u \to \infty$. On the basis of universality, however, we expect to obtain the same results even if we keep u finite.

The interaction K_{ij} depends only on the distance $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$ between the two lattice sites i and j, being given for example by

$$K_{ij} = K(\boldsymbol{r}_{ij}) = \begin{cases} K_0, & \text{if } i \text{ and } j \text{ are nearest neighbors;} \\ 0, & \text{otherwise.} \end{cases} \tag{6.73}$$

But we shall soon see that it is necessary to consider more general interaction matrices.

Let us now switch to Fourier space. We define $\boldsymbol{\phi}_k$ by the relation

$$\boldsymbol{\phi}_i = \frac{1}{N} \sum_k \boldsymbol{\phi}_k \, e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}, \qquad \forall i. \tag{6.74}$$

Using the relation

$$\sum_i \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i} = N\delta_{\boldsymbol{k},0}, \tag{6.75}$$

we obtain the following expression of $\boldsymbol{\phi}_{\boldsymbol{k}}$:

$$\phi_{\boldsymbol{k}} = \sum_i \boldsymbol{\phi}_i \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_i}. \tag{6.76}$$

Since the Hamiltonian (6.67) is quadratic in the $\boldsymbol{\phi}$s, it is easy to obtain its expression in Fourier space:

$$\mathcal{H}(\phi) = -\frac{1}{2N}\sum_{\boldsymbol{k}} K(\boldsymbol{k})\, \boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}} - \boldsymbol{h}\cdot\boldsymbol{\phi}_{\boldsymbol{k}=0}. \tag{6.77}$$

In this expression, $K(\boldsymbol{k})$ is the Fourier transform of the interaction matrix,

$$K(\boldsymbol{k}) = \sum_j K_{0j}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_j}, \tag{6.78}$$

and we take $\boldsymbol{k}$ in the first Brillouin zone. For a simple cubic lattice, as we saw in sec. 5.14, one has

$$K(\boldsymbol{k}) = 2K_0 \sum_{\alpha=1}^{d} \cos k_\alpha a_0 \simeq 2K_0 \left(d - \frac{k^2 a_0^2}{2} \right), \tag{6.79}$$

where a_0 is the lattice constant, and the first Brillouin zone is made by the wave vectors $\boldsymbol{k} = (k_\alpha)$ that satisfy

$$|k_\alpha| \leq \frac{\pi}{a_0} = \Lambda, \qquad \alpha = 1, \ldots, d. \tag{6.80}$$

We refer to $\Lambda = \pi/a_0$ as the **cutoff**. One should note that if the K_{ij} interaction is short-range, its Fourier transform $K(\boldsymbol{k})$ is analytic in $\boldsymbol{k}$, and therefore admits an expansion in powers of $\boldsymbol{k}$ around $\boldsymbol{k} = 0$. In the case of the simple cubic lattice, this expansion contains only even powers of $\boldsymbol{k}$. Thus it contains only k^2 to the second order (as we have just seen), while the k^4 order term will contain a combination of $(k^2)^2$ and of $\sum_\alpha (k_\alpha)^4$.

By introducing the weights $W(\boldsymbol{\phi}_i)$, we thus obtain the following expression of the partition function:

$$Z = \int \prod_{\boldsymbol{k}} \mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}}\, \mathcal{W}(\phi)\ \mathrm{e}^{-\mathcal{H}(\phi)}, \tag{6.81}$$

where $\mathcal{H}(\phi)$ is given by eq. (6.67) and $\mathcal{W}(\phi)$ is the product of all weights $W(\boldsymbol{\phi}_i)$ expressed in terms of the Fourier components $\boldsymbol{\phi}_{\boldsymbol{k}}$.

6.6.1 The Gaussian Model

The Gaussian model is obtained by considering a Gaussian weight $W(\boldsymbol{\phi}_i)$:

$$W(\phi_i) = \mathcal{N}\, \mathrm{e}^{-\lambda |\phi_i|^2/2}. \tag{6.82}$$

For $\lambda > 0$, one obtains a distribution centered around $\boldsymbol{\phi} = 0$. Then the functional $\mathcal{W}(\phi)\, \mathrm{e}^{-\mathcal{H}(\phi)}$ is Gaussian, and the partition function can be calculated explicitly. Setting $\boldsymbol{h} = 0$ for simplicity, we obtain

$$Z \propto \int \prod_{\boldsymbol{k}} \exp\left\{ -\frac{1}{2N} \sum_{\boldsymbol{k}} \left[\lambda + K(\boldsymbol{k})\right] \boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}} \right\}. \tag{6.83}$$

The value of the Fourier components $\boldsymbol{\phi}_{\boldsymbol{k}}$ depends on the number N of spins that make up the system, as one can see from eq. (6.76). In order to take the thermodynamic limit, it is convenient to redefine them so that their value becomes independent of N, which can be accomplished by rescaling them by a factor $1/\sqrt{N}$. The preceding expression then takes the form

$$Z \propto \int \prod_{\boldsymbol{k}} \mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}}\, \exp\left\{ -\frac{1}{2} \sum_{\boldsymbol{k}} \Delta(\boldsymbol{k}) \boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}} \right\}, \tag{6.84}$$

where

$$\Delta(\boldsymbol{k}) = r + ck^2 + \mathrm{O}\left(k^4\right), \tag{6.85}$$

in which

$$r = \lambda - 2dK_0, \tag{6.86a}$$

$$c = K_0 a_0^2. \tag{6.86b}$$

One can easily evaluate the next terms in the expansion in powers of $\boldsymbol{k}$. We can therefore integrate over the wave vectors $\boldsymbol{k}$ by employing the usual recipe:

$$\sum_{\boldsymbol{k}} \longrightarrow V \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^d} \equiv V \int_{\boldsymbol{k}}.$$

Thus we can express Z by $\int \prod_{\boldsymbol{k}} \mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}}\, \mathrm{e}^{-V\mathcal{H}(\phi)}$, where the Hamiltonian density (which we will henceforth denote again by $\mathcal{H}(\phi)$) is given by

$$\mathcal{H}(\phi) = \frac{1}{2} \int_{\boldsymbol{k}} \Delta(\boldsymbol{k})\, \boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}}. \tag{6.87}$$

The Gaussian model can be solved exactly. It is, however, more instructive to analyze it by means of the Kadanoff transformation in the following manner. First, we exploit universality so as to redefine the first Brillouin zone. In this fashion, we postulate that our field $\boldsymbol{\phi}_i$ has Fourier components $\boldsymbol{\phi}_{\boldsymbol{k}}$ for $\boldsymbol{k}$ included in a sphere in d dimensions of radius equal to the cutoff Λ. We then define the Kadanoff transformation as the sequential application of the following three operations.

1. We divide the field $\boldsymbol{\phi}$ into two components, $\boldsymbol{\phi}_{\boldsymbol{k}}^{<}$ and $\boldsymbol{\phi}_{\boldsymbol{k}}^{>}$, where $\boldsymbol{\phi}_{\boldsymbol{k}}^{<}$ has Fourier components for $|\boldsymbol{k}| < \Lambda/b$ (where $b > 1$ is the scaling factor), while $\boldsymbol{\phi}_{\boldsymbol{k}}^{>}$ has them for $\boldsymbol{k}$ belonging to the spherical shell $\Lambda/b < |k| < \Lambda$. We then integrate over $\boldsymbol{\phi}_{\boldsymbol{k}}^{>}$ in the expression for Z given by eq. (6.84). We thus obtain the following expression for the partition function:

$$Z = Z_0 \int \prod_{\boldsymbol{k}} \mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}}^{<} \, \mathrm{e}^{-V\mathcal{H}^{<}(\phi)}, \tag{6.88}$$

where Z_0 is a constant (which contributes to the free energy), and the Hamiltonian density $\mathcal{H}^{<}$ is given by

$$\mathcal{H}^{<}(\phi) = \frac{1}{2} \int_{\boldsymbol{k}}^{<} \Delta(\boldsymbol{k}) \, \boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}}, \tag{6.89}$$

where $\int_{\boldsymbol{k}}^{<}$ denotes the integrations over the range $|\boldsymbol{k}| < \Lambda/b$.

2. At this point, the maximum value of $|\boldsymbol{k}|$ is equal to $\Lambda/b = \pi/(a_0 b)$—this corresponds to a model in which the lattice constant is $a_0 b$ rather than a_0. In order to return to the preceding value of the step, we change variables in the expression of Hamiltonian density:

$$\boldsymbol{k} = \boldsymbol{k}'/b, \tag{6.90}$$

and we redefine $\boldsymbol{\phi}_{\boldsymbol{k}}^{<} = \bar{\boldsymbol{\phi}}_{\boldsymbol{k}'b}$. The Hamiltonian density becomes

$$\bar{\mathcal{H}} = \frac{1}{2} \int_{\boldsymbol{k}'} b^{-d} \, \Delta(\boldsymbol{k}'/b) \, \bar{\boldsymbol{\phi}}_{\boldsymbol{k}'} \cdot \bar{\boldsymbol{\phi}}_{-\boldsymbol{k}'} = \frac{1}{2} \int_{\boldsymbol{k}'} \bar{\Delta}(\boldsymbol{k}') \, \bar{\boldsymbol{\phi}}_{\boldsymbol{k}'} \cdot \bar{\boldsymbol{\phi}}_{-\boldsymbol{k}'}, \tag{6.91}$$

where

$$\bar{\Delta}(\boldsymbol{k}') = rb^{-d} + cb^{-(d+2)}k'^2 + \mathrm{O}\big(k'^4\big). \tag{6.92}$$

3. We now rescale the variables $\bar{\boldsymbol{\phi}}_{\boldsymbol{k}}$ by a factor ζ, whose value will be chosen appropriately later:

$$\bar{\boldsymbol{\phi}}_{\boldsymbol{k}} = \zeta \boldsymbol{\phi}_{\boldsymbol{k}}. \tag{6.93}$$

The final expression of the renormalized Hamiltonian density is therefore

$$\mathcal{H}' = \frac{1}{2} \int_{\boldsymbol{k}} \bar{\Delta}'(\boldsymbol{k}) \, \boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}}, \tag{6.94}$$

where

$$\Delta'(\boldsymbol{k}) = r' + c'k^2 + \mathrm{O}\big(k^4\big), \tag{6.95}$$

with

$$r' = rb^{-d}\zeta^2, \tag{6.96a}$$

$$c' = cb^{-(d+2)}\zeta^2. \tag{6.96b}$$

We can now choose the scaling factor ζ so as to have some fixed points. By choosing $\zeta = b^{d/2}$, one has $r' = r$, and therefore the point in which $r > 0$, $c = 0$ (and all the coefficients of k^n with $n > 2$ vanish) is a fixed point. It is easy to see that it is a stable fixed point. In fact, with this choice of ζ, one has $c' = cb^{-2} < c$, and c is therefore an irrelevant parameter.

Exercise 6.3 Show that all the coefficients of $k^n\,\phi_k\phi_{-k}$ with $n \neq 0$ are irrelevant parameters around this fixed point.

This fixed point corresponds to infinite temperature. Indeed, $\Delta(\boldsymbol{k})$ is proportional to the Fourier transform $K(\boldsymbol{k})$ of the interaction. At the fixed point, $\Delta(\boldsymbol{k})$, and thus $K(\boldsymbol{k})$, is constant, and therefore its antitransform K_{ij} is local—in other words, it vanishes whenever $i \neq j$. Thus, the system is made of noninteracting local variables $\boldsymbol{\phi}_i$—this is the situation one approaches by increasing the temperature in a system of variables interacting at short range. We will therefore call this fixed point ($r > 0$, $c = 0$, etc.) the **infinite temperature fixed point**.

We can otherwise choose $\zeta = b^{(d+2)/2}$. In this case, one has $c' = c$, but $r' = b^2 r$. In order to obtain the fixed point, we have to set $r = 0$, $c > 0$ and put all k^n coefficients with $n \geq 4$ to zero. It is then easy to see that r becomes the only relevant operator.

Exercise 6.4 Show that all the coefficients of $k^n\,\phi_k\phi_{-k}$ with $n \geq 4$ are irrelevant around this fixed point.

This fixed point corresponds to the critical point for the Gaussian model. In fact, let us calculate the correlation function

$$\left\langle \boldsymbol{\phi}_{k_0} \cdot \boldsymbol{\phi}_{-k_0} \right\rangle = \frac{1}{Z} \int \prod_{\boldsymbol{k}} \mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}} \left(\boldsymbol{\phi}_{k_0} \cdot \boldsymbol{\phi}_{-k_0} \right) \exp\left\{ -\frac{1}{2} \int_{\boldsymbol{k}'} \Delta(\boldsymbol{k}')\, \boldsymbol{\phi}_{\boldsymbol{k}'} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}'} \right\}. \tag{6.97}$$

The integral is Gaussian and factors into independent integrals, each relative to a pair $(\boldsymbol{k}, -\boldsymbol{k})$. These integrals cancel out with the analogous integrals in the denominator, except for the one relative to the pair $(\boldsymbol{k}_0, -\boldsymbol{k}_0)$, which gives

$$G(\boldsymbol{k}_0) = \left\langle \boldsymbol{\phi}_{k_0} \cdot \boldsymbol{\phi}_{-k_0} \right\rangle = \frac{1}{\Delta(\boldsymbol{k}_0)} \simeq \frac{1}{r + c k_0^2} \tag{6.98}$$

for small values of $|\boldsymbol{k}_0|$. The correlation function therefore has the Ornstein-Zernike form, and the coherence length ξ is proportional (for small values of r) to $r^{-1/2}$. It becomes infinite for $r = 0$, which therefore corresponds to the critical point. Assuming therefore $r \propto T - T_c$, we obtain the value of the exponent ν:

$$\nu = \frac{1}{2}. \tag{6.99}$$

The same result can be obtained from the recursion relations (6.96a):

$$\frac{1}{\nu} = \frac{1}{\ln b} \ln \left(\left. \frac{\mathrm{d}r'}{\mathrm{d}r} \right|_{r=r^*} \right) = 2. \tag{6.100}$$

It is not difficult to calculate the other exponents—in this fashion, one obtains, for example,

$$\alpha = 2 - \frac{d}{2}, \qquad \gamma = 1. \tag{6.101}$$

Since for $r = 0$, the correlation function is proportional to k^{-2}, one has $\eta = 0$. The exponent β is not defined in the Gaussian model. In fact, the region $T < T_c$ should be defined by $r < 0$. In this case, however, the partition function Z is not defined, because the integral over the $\boldsymbol{\phi}$ diverges. Analogously, the exponent δ is not defined. Let us note that all these exponents satisfy the scaling laws, and become equal to the classical exponents when $d = 4$.

Exercise 6.5 Obtain the value of the exponent α for $d < 4$ from the relation

$$C_h \propto |T - T_c|^{-\alpha},$$

where C_h is the thermal capacity at zero field, given by

$$C_h = -T \left.\frac{\partial^2 F}{\partial T^2}\right)_{h=0},$$

in which $F = -k_B T \ln Z$ is the Helmholtz free energy.

6.6.2 The ϕ^4 Model

Let us now express the weight $\mathcal{W}(\phi) = \prod_i W(\boldsymbol{\phi}_i)$, where $W(\boldsymbol{\phi}_i)$ is defined by eq. (6.72), in terms of the Fourier components $\boldsymbol{\phi}_{\boldsymbol{k}}$. We obtain

$$\mathcal{W}(\phi) \propto \exp\left\{ \frac{\lambda}{2N} \sum_{\boldsymbol{k}} \boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}} - \frac{u}{N^2} \sum_{\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4} \delta_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4,0} \left(\boldsymbol{\phi}_{\boldsymbol{k}_1} \cdot \boldsymbol{\phi}_{\boldsymbol{k}_2}\right) \left(\boldsymbol{\phi}_{\boldsymbol{k}_3} \cdot \boldsymbol{\phi}_{\boldsymbol{k}_4}\right) \right\}, \tag{6.102}$$

where λ is a coefficient that we do not need to make explicit, and where the Kronecker delta is required by translation invariance. In the infinite-volume limit, the sums over $\boldsymbol{k}$ become integrals over the d-dimensional sphere of radius Λ, the Kronecker delta becomes a Dirac delta, and we obtain the following expression of the partition function:

$$Z \propto \int \prod_{\boldsymbol{k}} \mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}}\, \mathrm{e}^{-V\mathcal{H}(\phi)}, \tag{6.103}$$

where the Hamiltonian is given by

$$\mathcal{H}(\phi) = \frac{1}{2} \int_{\boldsymbol{k}} \Delta(\boldsymbol{k})\, \boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}} + u \int_{\boldsymbol{k}_1} \int_{\boldsymbol{k}_2} \int_{\boldsymbol{k}_3} \int_{\boldsymbol{k}_4} \delta\left(\sum_{\alpha=1}^{4} \boldsymbol{k}\right) \left(\boldsymbol{\phi}_{\boldsymbol{k}_1} \cdot \boldsymbol{\phi}_{\boldsymbol{k}_2}\right) \left(\boldsymbol{\phi}_{\boldsymbol{k}_3} \cdot \boldsymbol{\phi}_{\boldsymbol{k}_4}\right) + \cdots, \tag{6.104}$$

where we have redefined the fields $\boldsymbol{\phi}$ and the coefficient u and understood the higher-order terms in the $\boldsymbol{\phi}$ variables. The model defined in this way is known in field theory as the ϕ^4 model. We can obtain the same model as a starting point from the Einstein theory of fluctuations, assuming the Landau expression for the availability. The idea is to consider

only those fluctuations whose wavelength is greater than a given elementary length, which we denote by a_0. The corresponding availability should be a regular function of the $\boldsymbol{\phi}_k$, since it results from integration over local degrees of freedom, and therefore Landau's expression can legitimately be assumed.

The model described by this Hamiltonian is no longer exactly solvable, and it is necessary to resort to some approximations. Let us rewrite the Hamiltonian in the following form:

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1, \tag{6.105}$$

where

$$\mathcal{H}_0 = \frac{1}{2}\int_k \Delta(\boldsymbol{k})\,\boldsymbol{\phi}_{\boldsymbol{k}}\cdot\boldsymbol{\phi}_{-\boldsymbol{k}}, \tag{6.106}$$

$$\mathcal{H}_{\mathrm{I}} = u\int_{k_1}\int_{k_2}\int_{k_3}\int_{k_4}\delta\left(\sum_{\alpha=1}^{4}\boldsymbol{k}\right)\left(\boldsymbol{\phi}_{\boldsymbol{k}_1}\cdot\boldsymbol{\phi}_{\boldsymbol{k}_2}\right)\left(\boldsymbol{\phi}_{\boldsymbol{k}_3}\cdot\boldsymbol{\phi}_{\boldsymbol{k}_4}\right)+\cdots \tag{6.107}$$

Let us now consider step 1 of the Kadanoff transformation. We have

$$\mathrm{e}^{-\mathcal{H}'}\zeta_0(\mathcal{H}) = \int\prod_{\Lambda/b<|\boldsymbol{k}|<\Lambda}\mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}}^{>}\,\mathrm{e}^{-(\mathcal{H}_0+\mathcal{H}_1)}. \tag{6.108}$$

We are not able to evaluate this integral exactly. We can, however, expand $\mathrm{e}^{-\mathcal{H}_1}$ in the form

$$\mathrm{e}^{-\mathcal{H}_1} = 1-\mathcal{H}_1+\frac{1}{2}\mathcal{H}_1^2+\cdots. \tag{6.109}$$

By substituting this expression in eq. (6.108), we obtain

$$\mathrm{e}^{-\mathcal{H}'} = Z_0\exp\left[-\mathcal{H}_0^{<}-\langle\mathcal{H}_1\rangle_0+\frac{1}{2}\left(\langle\mathcal{H}_1^2\rangle_0-\langle\mathcal{H}_1\rangle_0^2\right)+\cdots\right], \tag{6.110}$$

where, for any A, one has

$$\langle A\rangle_0 = \frac{1}{Z_0}\left[\int\prod_{\Lambda/b<|\boldsymbol{k}|<\Lambda}\mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}}^{>}\,\mathrm{e}^{-\mathcal{H}_0}A\right], \tag{6.111}$$

and Z_0 is a normalization factor. One therefore has

$$\mathcal{H}^{<} = \mathcal{H}_0^{<}+\langle\mathcal{H}_1\rangle_0-\frac{1}{2}\left(\langle\mathcal{H}_1^2\rangle_0-\langle\mathcal{H}_1\rangle_0^2\right)+\cdots. \tag{6.112}$$

This expression is a **cumulant expansion** of the renormalized Hamiltonian $\mathcal{H}^{<}$. At zero order in $\mathcal{H}_1$, we obtain again the quadratic Hamiltonian that we saw in the preceding section.

Let us consider the first nontrivial term, proportional to $\langle\mathcal{H}_1\rangle_0$. It is given by

$$-u\left(\frac{1}{Z_0}\right)\int_{k_1}\int_{k_2}\int_{k_3}\int_{k_4}\delta\left(\sum_{\alpha=1}^{4}\boldsymbol{k}_\alpha\right)\int\prod_{\boldsymbol{k}}\mathrm{d}\boldsymbol{\phi}_{\boldsymbol{k}}^{>}\,\mathrm{e}^{-V\mathcal{H}_0}\left(\boldsymbol{\phi}_{\boldsymbol{k}_1}\cdot\boldsymbol{\phi}_{\boldsymbol{k}_2}\right)\left(\boldsymbol{\phi}_{\boldsymbol{k}_3}\cdot\boldsymbol{\phi}_{\boldsymbol{k}_4}\right).$$

The terms that contribute to this expression can be sorted into the following classes.

1. Terms in which all the wave vectors $\boldsymbol{k}_\alpha$, $\alpha = 1, \ldots, 4$, have a magnitude greater than Λ/b. The contribution of these wave vectors is a constant that modifies the free energy, but not the interaction itself.

2. Terms in which only one of the wave vectors or three of them have a magnitude greater than Λ/b: these terms vanish because of symmetry when one integrates over $\boldsymbol{\phi}_k^>$.

3. Terms in which two vectors are larger than Λ/b. We can distinguish between two cases:

   - $\boldsymbol{k}_1 = -\boldsymbol{k}_2$, $\boldsymbol{k}_3 = -\boldsymbol{k}_4$. Let us assume that $|\boldsymbol{k}_3| > \Lambda/b$. We obtain the contribution

     $$n\left(\boldsymbol{\phi}_{k_1} \cdot \boldsymbol{\phi}_{-k_1}\right) \int_k^> G(\boldsymbol{k}),$$

     where $G(\boldsymbol{k})$ is the correlation function of the Gaussian model,

     $$G(\boldsymbol{k}) = \frac{1}{\Delta(\boldsymbol{k})}. \tag{6.113}$$

     The factor n is the result of the sum over the n components of the $\boldsymbol{\phi}$ fields. There are two terms of this type. We have denoted the integral extended to the spherical shell $\Lambda/b < |\boldsymbol{k}| < \Lambda$ by $\int_k^>$.

   - $\boldsymbol{k}_1 = -\boldsymbol{k}_3$, $\boldsymbol{k}_2 = -\boldsymbol{k}_4$ (or $\boldsymbol{k}_1 = -\boldsymbol{k}_4$, $\boldsymbol{k}_2 = -\boldsymbol{k}_3$). If we choose the first possibility and assume that $|\boldsymbol{k}_2| > \Lambda/b$, we obtain a contribution of the form

     $$\left(\boldsymbol{\phi}_{k_1} \cdot \boldsymbol{\phi}_{-k_1}\right) \int_k^> G(\boldsymbol{k}).$$

     There are four terms of this type.

4. Terms whose wave vectors are smaller than Λ/b in magnitude: these are not modified.

Summing up, we have

$$\begin{aligned}\langle \mathcal{H}_1 \rangle_0 = \text{const.} &+ u \int_{k_1}^< (2n+4) \left(\boldsymbol{\phi}_{k_1} \cdot \boldsymbol{\phi}_{-k_1}\right) \int_k^> G(\boldsymbol{k}) \\ &+ u \int_{k_1}^< \int_{k_2}^< \int_{k_3}^< \int_{k_4}^< \delta\left(\sum_{\alpha=1}^4 \boldsymbol{k}_\alpha\right) \left(\boldsymbol{\phi}_{k_1} \cdot \boldsymbol{\phi}_{k_2}\right) \left(\boldsymbol{\phi}_{k_3} \cdot \boldsymbol{\phi}_{k_4}\right).\end{aligned} \tag{6.114}$$

Remember that $\int_k^<$ refers to the integrals extended to the new Brillouin zone $|\boldsymbol{k}| < \Lambda/b$, while $\int_k^>$ refers to those extended to the spherical shell $\Lambda/b < |\boldsymbol{k}| < \Lambda$.

Starting from this expression, we perform steps 2 and 3 of the Kadanoff transformation. After some further reordering, we obtain the recurrence formulas:

$$r' = b^{-d}\zeta^2 \left[r + 2(2n+4)u \int_k^> G(\boldsymbol{k}) \right], \tag{6.115a}$$

$$c' = c b^{-(d+2)} \zeta^2, \tag{6.115b}$$

$$u' = u b^{-3d} \zeta^4. \tag{6.115c}$$

Let us consider eq. (6.115c). It implies that $u = 0$, $r = 0$ is a fixed point in the transformation: this is the Gaussian fixed point we studied in the preceding section. If we choose $\zeta = b^{(d+2)/2}$, we obtain $u' = ub^{4-d}$. Therefore, if $d > 4$, $u' < u$, and u is in this case an irrelevant perturbation with respect to the Gaussian fixed point. On the other hand, if $d < 4$, one has $u' > u$, and u is a relevant perturbation. We can therefore expect that the Gaussian model provides a good description of the critical behavior (at least above T_c) for $d > 4$.

In order to study the behavior of the system for $d < 4$, we must evaluate the transformations (6.115) and especially the value of u' at the next order in u. Actually, as we will see immediately, it will be sufficient to calculate the recurrence formula for the coefficient u only, which will depend only on the term $\langle \mathcal{H}_1^2 \rangle_0 - \langle \mathcal{H}_1 \rangle_0^2$. In this fashion, we obtain an expression of the form

$$u' = ub^{4-d} - u^2 w(b, r, n), \tag{6.116}$$

which implies the existence of a fixed point value u^* given by

$$u^* = \frac{1 - b^{4-d}}{w(b, r^*, n)}, \tag{6.117}$$

where r^* is the value of r corresponding to the critical fixed point. Obviously, our approximations will only be acceptable if u^* is small. Let us introduce the quantity $\epsilon = 4 - d$ and assume that we can treat it as a small quantity. We obtain

$$u^* \simeq \epsilon \frac{\ln b}{w(b, r^*, n)}. \tag{6.118}$$

It is therefore reasonable to calculate the renormalization transformations by means of perturbation techniques in the context of an expansion in ϵ.

6.6.3 Critical Exponents at the First Order in ϵ

We must therefore evaluate the expression

$$\langle \mathcal{H}_1^2 \rangle_0 - \langle \mathcal{H}_1 \rangle_0^2 .$$

The first term contains a product of eight $\boldsymbol{\phi}$ fields. Let us write $\boldsymbol{\phi} = \boldsymbol{\phi}^> + \boldsymbol{\phi}^<$, and evaluate the average over $\boldsymbol{\phi}^>$ with respect to the distribution $e^{-V\mathcal{H}_0}$. The terms with an odd number of $\boldsymbol{\phi}^>$ vanish because of symmetry. Let us consider the other terms one by one.

1. The terms with eight $\boldsymbol{\phi}^>$ factors are constant and contribute to free energy, but not to $\mathcal{H}^<$.

2. The terms with six $\boldsymbol{\phi}^>$ factors are proportional to $(\boldsymbol{\phi}^<)^2$ and therefore contribute to the renormalization of r and c. Since they are proportional to u^2, they are of an order higher than the ones we already evaluated, and we can disregard them for the time being.

3. The terms with four $\boldsymbol{\phi}^>$ factors are of the form $\phi^<_{k_1} \cdots \phi^<_{k_4} \langle \phi^>_{k_5} \cdots \phi^>_{k_8} \rangle$. Since the average is evaluated with a Gaussian distribution, it splits into the sum of the products of two averages of the $\langle \boldsymbol{\phi}^>_{\boldsymbol{k}} \cdot \boldsymbol{\phi}^>_{\boldsymbol{k}'} \rangle_0$ type. We thus obtain different terms, which can be sorted into two categories:

- Those in which the $\boldsymbol{\phi}^>$ factors that appear in each of these averages belong to the same factor $\mathcal{H}_1$. These terms contribute to $\langle \mathcal{H}_1 \rangle_0^2$ and are therefore cancelled out by the second term of our expression.

- Those in which the $\boldsymbol{\phi}^>$ factors in each of these two averages belong to two different factors $\mathcal{H}_1$. These are the contributions we must take into account. They are of the form

$$\phi^<_{k_1}\phi^<_{k_2}\phi^<_{k_3}\phi^<_{k_4}\int_{q_1}^<\int_{q_2}^<\int_{q_3}^<\int_{q_4}^<\delta\left(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{q}_1+\boldsymbol{q}_3\right)\delta\left(\boldsymbol{k}_2+\boldsymbol{k}_4+\boldsymbol{q}_3+\boldsymbol{q}_4\right)$$
$$\times\left\langle\boldsymbol{\phi}^>_{q_1}\cdot\boldsymbol{\phi}^>_{q_2}\right\rangle_0\left\langle\boldsymbol{\phi}^>_{q_3}\cdot\boldsymbol{\phi}^>_{q_4}\right\rangle_0.$$

The coefficient of $\phi^<_{k_1}\cdots\phi^<_{k_4}$ in this expression depends on the wave numbers $\boldsymbol{k}$, which means that the resulting renormalized interaction will not be purely local. As long as it is short-range, however (and therefore analytic as a function of $\boldsymbol{k}$), we can consider expanding it in series around $\boldsymbol{k}=0$. As we will see shortly, the terms of this expansion proportional to k^2 and so on are irrelevant and can be disregarded (at least at this order). We can therefore evalulate this coefficient for $\boldsymbol{k}_\alpha=0$. We obtain

$$\int_{q_1}^>\int_{q_2}^>\left\langle\boldsymbol{\phi}^>_{q_1}\cdot\boldsymbol{\phi}^>_{q_2}\right\rangle_0\left\langle\boldsymbol{\phi}^>_{-q_1}\cdot\boldsymbol{\phi}^>_{-q_2}\right\rangle_0.$$

Now translation invariance imposes that $\left\langle\boldsymbol{\phi}_{q_1}\cdot\boldsymbol{\phi}_{q_2}\right\rangle_0$ be proportional to $\delta(\boldsymbol{q}_1+\boldsymbol{q}_2)$, and this expression boils therefore down to

$$\int_q^>\left\langle\boldsymbol{\phi}_q\cdot\boldsymbol{\phi}_{-q}\right\rangle_0^2=\int_q^>G(\boldsymbol{q})^2.$$

We now have to calculate how many terms of this type are present, accounting for the sums over the different components of $\boldsymbol{\phi}$, among other things. This calculation is slightly technical and is discussed in section 6.6.4. The result is $8(n+8)$, which, if we take into account the factor $1/2$ that appears in the cumulant expansion, gives us a coefficient $4(n+8)$.

4. The terms with two $\boldsymbol{\phi}^>$ factors provide a contribution to the renormalized interaction, proportional to ϕ^6. Therefore, although our initial interaction contained only local terms in ϕ^4, we have obtained an interaction containing nonlocal terms (but short-range ones) in ϕ^4, as well as terms proportional to higher powers of $\boldsymbol{\phi}$. We will, however, be able to see that these terms are irrelevant and can be disregarded, at least at the lowest order in ϵ.

5. The terms with no $\boldsymbol{\phi}^>$ factors are contributions to $\langle \mathcal{H}_1 \rangle_0^2$ and are cancelled out by the second term of our expression.

Taking these new contributions into account, and applying the other two steps of the Kadanoff transformations, we obtain the recurrence formulas:

$$r'=b^{-d}\zeta^2\left[r+4(n+2)u\int_q^>G(\boldsymbol{q})\right], \tag{6.119a}$$

$$c'=cb^{-(d+2)}\zeta^2, \tag{6.119b}$$

$$u'=b^{-3d}\zeta^4\left[u-4(n+8)u^2\int_q^>G(\boldsymbol{q})^2\right]. \tag{6.119c}$$

Looking for a fixed point with a nonzero value of c, we see that we must set $\zeta = b^{(d+2)/2}$, as in the Gaussian model. Let us now have a look at the interaction terms proportional to ϕ^p, with $p > 4$. The relevant recurrence formula gives

$$v_p' = v_p b^{-(p-1)d} \zeta^p \left[1 + \mathrm{O}(u^2)\right]. \tag{6.120}$$

With our choice of ζ, we obtain $v_p' \simeq b^{p(-d/2+1)+d} v_p$, and therefore, v_p is irrelevant if $d \simeq 4$ and u is small at the fixed point. Analogously, if we consider some four-field interaction terms, which are proportional to k^p (with $p = 2, 4, \ldots$), we see that the corresponding coefficients are rescaled by a factor b^{4-d-p}, up to contributions of the order of u^2. These fields are therefore irrelevant, as long as ϵ and u are small.

We can now suppose $b = 1 + \delta\ell$, with $\delta\ell$ very small, and evaluate the wavenumbers in units of Λ. This allows us to evaluate the integrals that appear in these expressions in a very simple fashion:

$$\int_q^> G(q) \simeq K_d \frac{\delta\ell}{r + c}, \tag{6.121a}$$

$$\int_q^> G(q)^2 \simeq K_d \frac{\delta\ell}{(r + c)^2}, \tag{6.121b}$$

where $K_d = S_d/(2\pi)^d$ (S_d is the area of the unit sphere in d dimensions) is a factor that comes from the integration, and which we can get rid of by appropriately redefining u. With $\zeta = b^{(d+2)/2}$ (as we shall set from now on), we obtain

$$r' = (1 + \delta\ell)^2 \left[r + 4(n+2) \frac{u}{r + c} \delta\ell \right], \tag{6.122a}$$

$$c' = c, \tag{6.122b}$$

$$u' = (1 + \delta\ell)^\epsilon \left[u - 4(n+8) \frac{u^2}{(r + c)^2} \delta\ell \right]. \tag{6.122c}$$

We can therefore set $c = 1$ and write these relations as differential equations:

$$\frac{\mathrm{d}r}{\mathrm{d}\ell} = 2r + 4(n+2) \frac{u}{1 + r}, \tag{6.123a}$$

$$\frac{\mathrm{d}u}{\mathrm{d}\ell} = \epsilon u - 4(n+8) \frac{u^2}{(1 + r)^2}. \tag{6.123b}$$

The flow corresponding to these differential equations is shown in figure 6.7 in the two cases $d > 4$ (when the Gaussian fixed point is stable) and $d < 4$. In the latter case, the Gaussian fixed point becomes unstable, and a new stable fixed point appears with $u^* > 0$. Always for $c = 1$, this fixed point (r^*, u^*) is given (at first order in ϵ) by

$$r^* = -\frac{1}{2} \frac{n+2}{n+8} \epsilon, \tag{6.124a}$$

$$u^* = \frac{\epsilon}{4(n+8)}. \tag{6.124b}$$

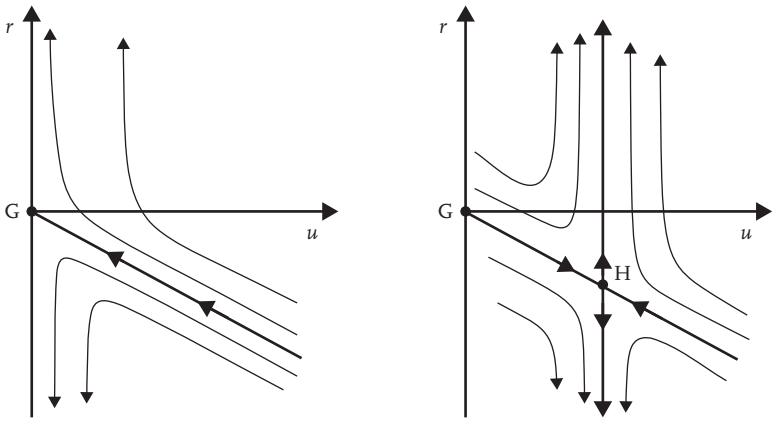


FIGURE 6.7. Renormalization flow in the (u, r) plane for the ϕ^4 model. Left: $d > 4$. The G (Gaussian) fixed point is stable for the entire critical line $r = r_c(u) = -2(n+2)u$. Right: $d < 4$. The G fixed point becomes unstable, and the H (Heisenberg) fixed point appears, which is stable for points on the critical line for which $u > 0$.

This fixed point is an attractor for Hamiltonians starting on the critical line $r = r_c(u) = -2(n+2)u$. It describes therefore the critical behavior of the n-component Heisenberg model for $d < 4$. We will therefore call it the **Heisenberg fixed point**.

Let us now consider the linearized flow equations around this fixed point:

$$\frac{\mathrm{d}}{\mathrm{d}\ell}\begin{pmatrix} r \\ u \end{pmatrix} = \begin{pmatrix} 2 - [(n+2)/(n+8)]\epsilon, & 4(n+2) \\ 0, & -\epsilon \end{pmatrix}\begin{pmatrix} r - r^* \\ u - u^* \end{pmatrix}. \tag{6.125}$$

The eigenvalues of the matrix that defines the linearized flow are

$$\lambda_r = 2 - \frac{n+2}{n+8}\epsilon, \tag{6.126a}$$

$$\lambda_u = -\epsilon. \tag{6.126b}$$

The first eigenvalue determines the value of the exponent ν:

$$\nu = \frac{1}{\lambda_r} = \frac{1}{2} + \frac{n+2}{4(n+8)}\epsilon + \mathrm{O}(\epsilon^2). \tag{6.127}$$

The second eigenvalue controls the approach to the fixed point.

The second independent critical exponent, η, can be obtained from the expression of ζ. When the lengths are rescaled by a factor $1/b$ at the fixed point, the correlation function in real space is rescaled by a factor ζ^2. Therefore,

$$G(\boldsymbol{r}/b) = \zeta^2 G(\boldsymbol{r}) = b^{d+2} G(\boldsymbol{r}). \tag{6.128}$$

This implies $G(\boldsymbol{r}) \sim |\boldsymbol{r}|^{-(d+2)}$, which should be compared with the relation $G(\boldsymbol{r}) \sim |\boldsymbol{r}|^{-(d+2-\eta)}$ that defines the exponent η. We thus obtain

$$\eta = \mathrm{O}(\epsilon^2). \tag{6.129}$$

The other exponents can be obtained from the scaling laws, stopping of course at the first order in ϵ. One thus obtains

$$\alpha = \frac{4-n}{2(n+8)}\epsilon + \mathrm{O}(\epsilon^2), \tag{6.130a}$$

$$\beta = \frac{1}{2} - \frac{3}{2(n+8)}\epsilon + \mathrm{O}(\epsilon^2), \tag{6.130b}$$

$$\gamma = 1 + \frac{n+2}{2(n+8)}\epsilon + \mathrm{O}(\epsilon^2), \tag{6.130c}$$

$$\delta = 3 + \epsilon + \mathrm{O}(\epsilon^2). \tag{6.130d}$$

By evaluating these expressions for $\epsilon = 1$, one does not obtain particularly good results—the interest in this approach lies in the fact that it can be systematically improved (calculations up to the fifth order in ϵ have been performed [Klei01]) and that it can provide general qualitative information, about the dependence on n, for instance. Most of all, this method allows us to systematically control universality properties in the most general fashion.

6.6.4 Multiplicity of the Contributions to Renormalization of u

In this section, I show how one obtains the coefficient $4(n+8)$ for the multiplicity in eq. (6.119c).

By writing all the sums, both over the vectors and over the field components, $\langle \mathcal{H}_1^2 \rangle_0$ can be written

$$\begin{aligned}\langle \mathcal{H}_1^2 \rangle_0 = \frac{1}{Z_0} \int \prod_{\boldsymbol{q}} \mathrm{d}\boldsymbol{\phi}_{\boldsymbol{q}}^{>}\, \mathrm{e}^{-V\mathcal{H}_0} \sum_{\alpha_1\alpha_2} \sum_{\beta_1\beta_2} \int_{k_1} \cdots \int_{k_4} \int_{q_1} \cdots \int_{q_4} \\ \times \delta\left(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 + \boldsymbol{k}_4\right) \delta\left(\boldsymbol{q}_1 + \boldsymbol{q}_2 + \boldsymbol{q}_3 + \boldsymbol{q}_4\right) \\ \times \left(\phi_{\boldsymbol{k}_1,\alpha_1} \phi_{\boldsymbol{k}_2,\alpha_1} \phi_{\boldsymbol{k}_3,\alpha_2} \phi_{\boldsymbol{k}_4,\alpha_2}\right) \left(\phi_{\boldsymbol{q}_1,\beta_1} \phi_{\boldsymbol{q}_2,\beta_1} \phi_{\boldsymbol{q}_3,\beta_2} \phi_{\boldsymbol{q}_4,\beta_2}\right).\end{aligned} \tag{6.131}$$

The terms we are considering are those in which two fields $\boldsymbol{\phi}^{>}$ belonging to the first factor (denoted by $(\boldsymbol{k}, \alpha)$) are paired each with a field belonging to the second factor (denoted by $(\boldsymbol{q}, \beta)$). One thus obtains two factors $\left\langle \phi_{\boldsymbol{k},\alpha}^{>} \phi_{\boldsymbol{q},\beta}^{>} \right\rangle_0$ that vanish, unless one has $\boldsymbol{k} = -\boldsymbol{q}$ and $\alpha = \beta$.

The terms one obtains are of three types.

1. Terms of the type

   $$\langle \phi_{\boldsymbol{k}_1,\alpha_1}^{>} \phi_{\boldsymbol{q}_1,\beta_1}^{>} \rangle_0 \langle \phi_{\boldsymbol{k}_2,\alpha_1}^{>} \phi_{\boldsymbol{q}_2,\beta_1}^{>} \rangle_0 \phi_{\boldsymbol{k}_3,\alpha_2}^{<} \phi_{\boldsymbol{k}_4,\alpha_2}^{<} \phi_{\boldsymbol{q}_3,\beta_2}^{<} \phi_{\boldsymbol{q}_4,\beta_2}^{<}.$$

   In this case, one can freely sum over $\alpha_1 = \beta_1$, and one therefore obtains a contribution that is proportional to the number n of components of the $\boldsymbol{\phi}^{>}$ field. It is easy to see that there are eight contributions of this sort: one can choose α_1 or α_2 and connect it with β_1 or β_2, and there are two ways to make the connection.

2. Terms of the type

   $$\langle \phi_{\boldsymbol{k}_1,\alpha_1}^{>} \phi_{\boldsymbol{q}_1,\beta_1}^{>} \rangle_0 \langle \phi_{\boldsymbol{k}_2,\alpha_2}^{>} \phi_{\boldsymbol{q}_3,\beta_2}^{>} \rangle_0 \phi_{\boldsymbol{k}_3,\alpha_1}^{<} \phi_{\boldsymbol{k}_4,\alpha_2}^{<} \phi_{\boldsymbol{q}_2,\beta_1}^{<} \phi_{\boldsymbol{q}_4,\beta_2}^{<}.$$

In this case, one must have $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$, and there are therefore no free components of the $\boldsymbol{\phi}^>$ field over which to sum. The result therefore does not depend on n. There are 32 terms of this type: one can connect any sort of α field with any of the four β fields (16 choices), and one then needs to connect an α field that contains a different component to one of the β fields with a component that differs from the one previously connected (four choices). One then needs to divide by 2 because the same result can be obtained by making one's choices in the two α fields in the reverse order.

3. Terms of the type

$$\langle \phi^>_{\boldsymbol{k}_1,\alpha_1} \phi^>_{\boldsymbol{q}_1,\beta_1} \rangle_0 \langle \phi^>_{\boldsymbol{k}_3,\alpha_2} \phi^>_{\boldsymbol{q}_2,\beta_1} \rangle_0 \phi^<_{\boldsymbol{k}_2,\alpha_1} \phi^<_{\boldsymbol{k}_4,\alpha_2} \phi^<_{\boldsymbol{q}_3,\beta_2} \phi^<_{\boldsymbol{q}_4,\beta_2}.$$

In this case, one must have $\alpha_1 = \beta_1 = \alpha_2$, and there are therefore once again no free components of the $\boldsymbol{\phi}^>$ field. Using the same type of reasoning as earlier, we see that there are 32 components of this type.

The multiplicity of contributions is therefore $8n + 64$. One can check this result by setting $n = 1$. One obtains a multiplicity equal to 72. On the other hand, for $n = 1$, one has $(4 \times 4 \times 3 \times 3)/2 = 72$ terms (four possible αs with four possible βs, then one of the remaining three αs with one of the remaining βs, divided by 2, because one obtains the same result by switching the two pairs). Taking the factor $1/2$ that comes from the cumulant expansion into account, we obtain the coefficient $4(n + 8)$ in eq. (6.119c).

These evaluations are greatly facilitated if diagrammatic techniques are employed, and these are presented in almost all the books on the topic.

6.6.5 $1/n$ Expansion of Critical Exponents

An alternative approach takes advantage of the exact solution of the $n \to \infty$ limit, seen in sec. 5.19, to set up a systematic expansion of critical properties in powers of $1/n$ (see, e.g., [Abe73], [Ma73], and [Ma76, sec. IX, 4]). The gist is to evaluate the logarithmic deviations from the $n \to \infty$ behavior by a suitably adapted perturbation theory: for example, the inverse correlation function at T_c behaves as

$$C^{-1}(\boldsymbol{k}, T_c) \sim k^{2-\eta} = k^2 \big(1 + \eta \ln k + \mathrm{O}(\eta^2)\big), \tag{6.132}$$

where we have assumed that η is small. Therefore, if $\eta \propto 1/n$, it could be obtained by evaluating $C^{-1}(\boldsymbol{k}, T_c)$ perturbatively to first order in $1/n$. I refer you to the cited literature for the details.

Exercise 6.6 (Critical isotherm) Discuss the behavior of the $n \to \infty$ model for $h \neq 0$, and obtain the value of exponent δ directly. Compare this with the result obtained from the scaling laws:

$$\delta = \frac{d+2}{d-2}; \qquad 2 < d \leq 4.$$

6.7 Quadratic Anisotropy and Crossover

Let us now consider a Hamiltonian in which the different components of the order parameter interact with different intensities. If the order parameter $\boldsymbol{\phi}$ is an n-component unit vector, for instance, we can consider the following generalization of Heisenberg's Hamiltonian:

$$\mathcal{H} = -\frac{1}{2}\sum_{ij} K_{ij} \sum_{\alpha=1}^{n} (1+\Delta_\alpha)\, \phi_i^\alpha \phi_j^\alpha. \tag{6.133}$$

Let us assume that $\Delta_\alpha = \Delta > 0$ for $\alpha = 1, \ldots, m$ and $\Delta_\alpha = 0$ for $\alpha = m+1, \ldots, n$. In this case, the first m components of $\boldsymbol{\phi}$ will interact more strongly than the others, and we can expect the symmetry breaking to privilege them: the critical behavior of the model will be the one of a model with an m-dimensional order parameter. If Δ is very small, on the other hand, we expect that the system, for T not too close to T_c (or for $T = T_c$, but for fairly small distances), to behave approximately like the symmetrical model.

Let us see how one can deal with the effect of this type of perturbation at the first order in ϵ. It is easier to see it directly in Fourier space. We then see that Δ modifies the quadratic part of the Hamiltonian density. A term of this type is therefore called a **quadratic anisotropy** term. Let us attempt to describe its effect in the simplest fashion. Let us therefore consider a Hamiltonian density of the form

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1, \tag{6.134}$$

where $\mathcal{H}_1$ has the form seen in the preceding section, while

$$\mathcal{H}_0 = \frac{1}{2}\int_{\boldsymbol{k}} \sum_{\alpha=1}^{n} \left(r_\alpha + k^2\right) \phi_{\boldsymbol{k}}^\alpha \phi_{-\boldsymbol{k}}^\alpha. \tag{6.135}$$

By going through the calculations sketched in the previous section, we obtain the recurrence relation for the coefficients r_α:

$$\frac{\mathrm{d}r_\alpha}{\mathrm{d}\ell} = 2r_\alpha + 4u\left(\frac{2}{1+r_\alpha} + \sum_{\beta=1}^{n} \frac{1}{1+r_\beta}\right). \tag{6.136}$$

By linearizing around the Heisenberg fixed point and setting $\delta r_\alpha = (r_\alpha - r^*)$, $\delta r = \sum_\alpha \delta r_\alpha$, and $\Delta_\alpha = \delta r_\alpha - \delta r/n$, we obtain

$$\frac{\mathrm{d}\delta r}{\mathrm{d}\ell} = [2 - 4(n+2)u^*]\, \delta r, \tag{6.137a}$$

$$\frac{\mathrm{d}\Delta_\alpha}{\mathrm{d}\ell} = 2(1-4u^*)\Delta_\alpha. \tag{6.137b}$$

The first equation governs the behavior of the correlation length around the Heisenberg fixed point and shows that the exponent ν is not modified. The second equation determines the critical behavior of the system with a small anisotropy.

Let us, in fact, consider the system at the critical point (therefore, $\delta r = 0$) and with, for example, $\Delta_1 > 0$, $\Delta_\alpha = 0$, $\alpha = 2, \ldots, n$. By solving the recurrence relations, we obtain $\Delta_1(\ell) = \Delta(0)\, \mathrm{e}^{\psi \ell}$, where the exponent ψ is given by

$$\psi = 2(1 - 4u^*) = 1 - \frac{2\epsilon}{n+8} + \mathrm{O}(\epsilon^2). \tag{6.138}$$

Let us recall that the length rescaling factor b is given by $b = \mathrm{e}^{\ell}$. The correlation function $G(\boldsymbol{r}, \Delta)$, where $\Delta = \Delta_1(0)$, satisfies the relation

$$G(\boldsymbol{r}, \Delta) = b^{-(d+2)} G(\boldsymbol{r}/b, \Delta b^{\psi}). \tag{6.139}$$

By choosing $b = |\boldsymbol{r}|$ we obtain $G(\boldsymbol{r}, \Delta) = |\boldsymbol{r}|^{-(d+2)} f(\boldsymbol{r}/\Delta^{-1/\psi})$, where $f(x)$ is a certain function. We see, therefore, that there is a certain characteristic length $\ell_\Delta \sim \Delta^{-1/\psi}$ that governs the effects of anisotropy at the critical point. For $|\boldsymbol{r}| \gg \ell_\Delta$, the behavior is similar to that with $\Delta = 0$ (isotropic model), while for $|\boldsymbol{r}| \ll \ell_\Delta$, the behavior corresponds to that of large values of Δ, and therefore to a model with a smaller number of dimensions of the order parameter.

If we instead consider $T \neq T_\mathrm{c}$, we have to compare the correlation length $\xi \sim |T - T_\mathrm{c}|^{-\nu}$ with ℓ_Δ. For $\boldsymbol{k} = 0$, therefore, the change in behavior occurs for $|T - T_\mathrm{c}| \sim \ell^{-1/\varphi}$, where $\varphi = \psi\nu$ is called the **crossover exponent**:

$$\varphi = \psi\nu = 1 + \frac{n}{2(n+8)}\epsilon + \mathrm{O}(\epsilon^2). \tag{6.140}$$

More particularly, this result implies that $T_\mathrm{c}(\Delta) = T_\mathrm{c}(0) + \mathrm{O}\left(\Delta^{1/\varphi}\right)$. Since $\varphi > 1$, the line of critical points exhibits a cusp in $\Delta = 0$. In this way, the critical behavior of the model is described by one fixed point further from the critical point, and by a different one closer to it. This phenomenon is called **crossover**.

We can summarize this result by saying that the critical surface is dominated by three fixed points: the $O(n)$-symmetric one, the $O(n-1)$-symmetric one, and the Ising fixed point (corresponding to $n = 1$). If $\Delta > 0$, the renormalization flow eventually approaches the Ising-like fixed point. Thus the region $\Delta > 0$ of the critical surface is the attraction basin of this fixed point. When $\Delta < 0$, the relevant fixed point is the $O(n-1)$-symmetric one. The attraction basin of the $O(n)$-symmetric fixed point is the region with $\Delta = 0$, and has therefore one dimension less than that of the other fixed points.

6.8 Critical Crossover

The most common case of crossover is obviously that by means of which the behavior of a system is described by a Gaussian fixed point when it is far from the critical temperature, and by a Heisenberg fixed point close to the critical temperature—the transition from one behavior to the other occurs at a temperature determined by the Ginzburg criterion. In this section, I show how the eqs. (6.123) describing the renormalization flow at the first order in ϵ allow one to obtain the crossover behavior of the susceptibility.

Let us rewrite the eqs. (6.123) in the form

$$\frac{\mathrm{d}r}{\mathrm{d}\ell} = 2r + A\frac{u}{1+r}, \tag{6.141a}$$

$$\frac{\mathrm{d}u}{\mathrm{d}\ell} = \epsilon u - B\frac{u^2}{(1+r)^2}, \tag{6.141b}$$

where $A = 4(n+2)$ and $B = 4(n+8)$. Taking into account that the exponent η vanishes at this order in ϵ, the correlation function $C(\boldsymbol{k}, r, u) = \langle \phi_{\boldsymbol{k}} \cdot \phi_{-\boldsymbol{k}} \rangle$ satisfies the relation

$$C(\boldsymbol{k}, r_0, u_0) = \mathrm{e}^{2\ell} C\left(\mathrm{e}^{\ell}\boldsymbol{k}, r(\ell), u(\ell)\right), \tag{6.142}$$

where $[r(\ell), u(\ell)]$ is the solution of eqs. (6.141) that satisfies $[r(0), u(0)] = [r_0, u_0]$. We can use this relation to express $C(\boldsymbol{k}, r, u)$ for small values of r and $|\boldsymbol{k}|$ (in the proximity of the critical point) in terms of $C\left(\mathrm{e}^{\ell}\boldsymbol{k}, r(\ell), u(\ell)\right)$, where $r(\ell)$ and $\mathrm{e}^{\ell}|\boldsymbol{k}|$ are of order 1, which should allow us to evaluate with ease C in perturbation theory.

We will soon see that, at the first order in ϵ, we can set $r = 0$ in the flow equation for u, so that eq. (6.141b) takes the form

$$\frac{\mathrm{d}u}{\mathrm{d}\ell} = \epsilon u - Bu^2. \tag{6.143}$$

This equation can be solved exactly, and has the solution

$$u(\ell) = u_0 \frac{\mathrm{e}^{\epsilon\ell}}{Q(\ell)}, \tag{6.144}$$

where

$$Q(\ell) = 1 + \frac{Bu_0}{\epsilon}\left(\mathrm{e}^{\epsilon\ell} - 1\right). \tag{6.145}$$

It is clear that

$$u(\ell) \to \begin{cases} u_0, & \text{for } \ell \to 0, \\ \epsilon/B = u^*, & \text{for } \ell \to \infty. \end{cases} \tag{6.146}$$

The transition from one value to the other occurs for $(Bu_0/\epsilon)\mathrm{e}^{\epsilon\ell} \sim 1$. We will always assume that u_0 is of order ϵ.

Let us denote the solution we just found by $u_1(\ell)$ and the solution of eq. (6.141b) by $u(\ell)$. Let us set $u(\ell) = f(\ell)u_1(\ell)$. Then,

$$u_1\frac{\mathrm{d}f}{\mathrm{d}\ell} = -Bu_1^2\left(f^2 - f\right) + f^2u_1^2\left[\frac{1}{(1+r)^2} - 1\right]. \tag{6.147}$$

Since u_1 is of order ϵ, and $f \to 1$ for $\ell \to 0$, one has $f = 1 + \mathrm{O}(\epsilon)$: we can therefore set $u = u_1$ at the first order in ϵ.

The solution of eq. (6.141a) is a more delicate matter, especially since we must consider the possibility that r becomes of order 1. Let us introduce the quantity

$$t(\ell) = r(\ell) + \frac{1}{2}A\,u(\ell) - \frac{1}{2}A\,u(\ell)r(\ell)\ln[1 + r(\ell)], \tag{6.148}$$

and take into account the fact that $\mathrm{d}u/\mathrm{d}\ell = \mathrm{O}(\epsilon^2)$. By neglecting contributions of order ϵ^2, one then has

$$\begin{aligned}\frac{\mathrm{d}t}{\mathrm{d}\ell} &= \frac{\mathrm{d}r(\ell)}{\mathrm{d}\ell}\left[1 - \frac{1}{2}A\,u(\ell)\ln[1+r(\ell)] - \frac{1}{2}A\,u(\ell)\frac{r(\ell)}{1+r(\ell)}\right] \\ &= \left[2r(\ell) + A\frac{u(\ell)}{1+r(\ell)}\right]\left[1 - \frac{1}{2}A\,u(\ell)\ln[1+r(\ell)] - \frac{1}{2}A\,u(\ell)\frac{r(\ell)}{1+r(\ell)}\right] \\ &= 2r(\ell) + A\,u(\ell) - A\,u(\ell)r(\ell)\ln[1+r(\ell)] + A\,u(\ell)r(\ell) \\ &= [2 - A\,u(\ell)]t(\ell) + \mathrm{O}\left(\epsilon^2\right).\end{aligned} \tag{6.149}$$

Then

$$t(\ell) = \mathrm{e}^{S(\ell)}t(0), \tag{6.150}$$

where

$$S(\ell) = 2\ell - A\int_0^\ell \mathrm{d}\ell'\,u(\ell'). \tag{6.151}$$

We see that, if $t(0) = 0$, then $t(\ell) = 0$, $\forall\ell$. We can therefore identify t with the distance from the critical surface. The transition therefore occurs for $t(0) = 0$—in other words, for $r(0) = -(A/2)u(0) + \mathrm{O}(\epsilon^2)$.

Evaluating $S(\ell)$ explicitly, we obtain

$$\begin{aligned}S(\ell) &= 2\ell - Au_0\int_0^\ell \mathrm{d}\ell'\,\frac{\mathrm{e}^{\epsilon\ell'}}{1 + (Bu_0/\epsilon)\left(\mathrm{e}^{\epsilon\ell'} - 1\right)} \\ &= 2\ell - \frac{A}{B}\ln Q(\ell).\end{aligned} \tag{6.152}$$

Therefore

$$t(\ell) = \mathrm{e}^{2\ell}\,[Q(\ell)]^{-A/B}\,t(0), \tag{6.153}$$

which has the limit behavior

$$t(\ell) \simeq \begin{cases}\mathrm{e}^{2\ell}\,t(0), & \text{for } \ell \to 0, \\ \mathrm{e}^{(2-A\epsilon/B)\ell}\,(Bu_0/\epsilon)\,t(0), & \text{for } \ell \to \infty.\end{cases} \tag{6.154}$$

This is the crossover we are looking for. The mean-field behavior is valid as long as $(Bu_0/\epsilon)\mathrm{e}^{\epsilon\ell} \ll 1$. For $(Bu_0/\epsilon)\mathrm{e}^{\epsilon\ell} \gg 1$, the behavior is dictated by the Heisenberg fixed point, with $t(\ell) \simeq \mathrm{e}^{\ell/\nu}T(0)$, where $1/\nu = 2 - [(n+2)/(n+8)]\epsilon$. The crossover occurs at the temperature defined by $t = (Bu_0/\epsilon)^{2/\epsilon}$.

We must now relate the physical quantities to the solution obtained for $t(\ell)$. Let us choose ℓ^* such that

$$t(\ell^*) = 1. \tag{6.155}$$

At this point, it is possible to use perturbation theory to evaluate the correlation functions. The susceptibility χ is proportional to the correlation function for $\boldsymbol{k}=0$. In perturbation theory, it is more convenient to evaluate the *inverse* of $C(\boldsymbol{k}, r, u)$. In effect, at the lowest order in u, one has

$$C^{-1}(\boldsymbol{k}, r, u) = r + k^2 + Au \int_q G(\boldsymbol{q}, r), \tag{6.156}$$

where $G(\boldsymbol{q}, r) = 1/(r+q^2)$ is the correlation function of the Gaussian model. This relation is derived at the end of this section. In $d=4$, one has, at the lowest order in the perturbation,

$$\begin{aligned} C^{-1}(0, r, u) &= r + Au \int_q \frac{1}{r+q^2} = r + \frac{1}{2}Au - \frac{1}{2}Au\, r\,[\ln(1+r) - \ln r] \\ &\simeq t + \frac{1}{2} A\, u \ln t. \end{aligned} \tag{6.157}$$

Therefore, for $\ell = \ell^*$, one has $\chi(r(\ell^*), u(\ell^*)) = 1/t(\ell^*) = 1$. By using eq. (6.142), we obtain

$$\chi(r, u) = \mathrm{e}^{2\ell^*} \chi(r(\ell^*), u(\ell^*)) = \mathrm{e}^{2\ell^*}, \tag{6.158}$$

where ℓ^* satisfies the equation

$$\mathrm{e}^{2\ell^*} \left[1 + \frac{Bu_0}{\epsilon} \left(\mathrm{e}^{\epsilon\ell^*} - 1\right)\right]^{-A/B} t = 1, \tag{6.159}$$

in which we have set $t = t(0)$. An approximate solution of this equation is given by

$$\mathrm{e}^{\ell^*} = t^{-1/2} \left(1 + B\frac{u_0}{\epsilon t^{\epsilon/2}}\right)^{A/2B}. \tag{6.160}$$

We thus have

$$\chi = t^{-1} \left(1 + B\frac{u_0}{\epsilon t^{\epsilon/2}}\right)^{A/B} \sim \begin{cases} t^{-1}, & \text{for } (Bu_0/\epsilon) \ll t^{\epsilon/2}, \\ t^{-(1+\epsilon A/2B)}, & \text{for } (Bu_0/\epsilon) \gg t^{\epsilon/2}. \end{cases} \tag{6.161}$$

The crossover can be visualized by introducing the effective exponent γ_{eff}, given by

$$\gamma_{\text{eff}} = -\frac{\mathrm{d}\ln\chi}{\mathrm{d}\ln t} \simeq 1 + \frac{A}{2}\frac{u_0/\epsilon}{t^{\epsilon/2} + (Bu_0/\epsilon)}. \tag{6.162}$$

Figure 6.8 shows $\gamma_{\text{eff}}(t)$ evaluated with $n=1$, $\epsilon=1$, and $u_0 = 0.1$.

6.8.1 Derivation of Eq. (6.156)

It is convenient to start from the expression of the average of ϕ_i^α:

$$\left\langle \phi_i^\alpha \right\rangle = \frac{1}{Z} \operatorname{Tr} \phi_i^\alpha\, \mathrm{e}^{-\mathcal{H}}, \tag{6.163}$$

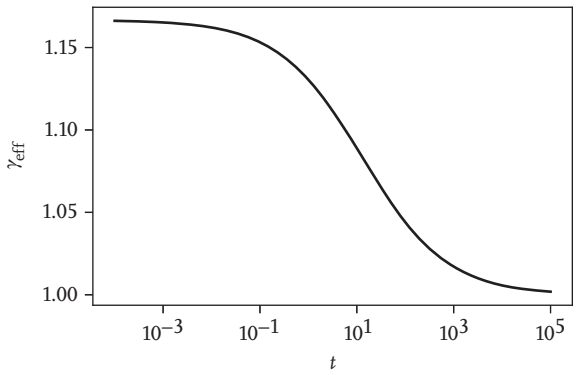


FIGURE 6.8. Effective exponent γ_{eff} as a function of $t \propto |T - T_c|$ for $n = 1$, $\epsilon = 1$, and $u_0 = 0.1$.

where $\mathrm{Tr} = \int \prod_{i\alpha} \mathrm{d}\phi_i^\alpha$,

$$Z = \mathrm{Tr}\, \mathrm{e}^{-\mathcal{H}}, \tag{6.164}$$

and $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$, with

$$\mathcal{H}_0 = \frac{1}{2} \sum_{ij} \sum_{\alpha} \phi_i^\alpha G_{ij}^{-1} \phi_j^\alpha, \tag{6.165a}$$

$$\mathcal{H}_1 = u \sum_i \left(\sum_\alpha \phi_i^\alpha \right)^2. \tag{6.165b}$$

Here G_{ij} is the correlation function of the Gaussian model. The integral of a total derivative vanishes when the integrand vanishes at infinity, as in our case. We thus have

$$\begin{aligned} 0 &= \mathrm{Tr}\, \frac{\mathrm{d}}{\mathrm{d}\phi_j^\alpha} \left(\phi_i^\alpha \, \mathrm{e}^{-\mathcal{H}} \right) = Z \left(\delta_{ij} - \mathrm{Tr}\, \phi_i^\alpha \frac{\partial \mathcal{H}}{\partial \phi_j^\alpha} \right) \\ &= Z \left[\delta_{ij} - \sum_k G_{jk}^{-1} \left\langle \phi_i^\alpha \phi_k^\alpha \right\rangle - 4u \left\langle \phi_i^\alpha \phi_j^\alpha \sum_\beta \left(\phi_j^\beta \right)^2 \right\rangle \right]. \end{aligned} \tag{6.166}$$

At the lowest order, we can set

$$\sum_\beta \left\langle \phi_i^\alpha \phi_j^\alpha \left(\phi_j^\beta \right)^2 \right\rangle = (n+2) \left\langle \phi_i^\alpha \phi_j^\alpha \right\rangle \left\langle \left(\phi_j^\beta \right)^2 \right\rangle \tag{6.167}$$

(where one no longer sums over the indexes) and evaluate the average of ϕ^2 in the Gaussian model. Multiplying by C_{jk}^{-1} and summing over j with the help of sec. 6.6.4, we obtain

$$C_{ik}^{-1} = G_{ik}^{-1} + 4u(n+2) \left\langle \left(\phi_j^\beta \right)^2 \right\rangle. \tag{6.168}$$

By passing to the Fourier transforms, we arrive at eq. (6.156).

6.9 Cubic Anisotropy

Perturbations that explicitly break the $O(n)$ of the order parameter give rise to an interesting structure of fixed points and flow over the critical surface. We describe these effects in the ϕ^4 model by introducing, beyond the usual $O(n)$-symmetric interaction term, a new interaction term of the form:

$$\mathcal{H}_{\text{anis}} = v \sum_i \sum_{\alpha=1}^{n} \phi_{i\alpha}^4. \tag{6.169}$$

For $u = 0$, the n-component model is reduced to a collection of n independent models with $n = 1$ (Ising). It is easy to see that for $d > 4$, the new term is irrelevant. By using the techniques we have developed, one obtains the first-order flow equations in ϵ:

$$\frac{\mathrm{d}r}{\mathrm{d}\ell} = 2r + 4\left[(n+2)u + 3v\right]\frac{1}{1+r}, \tag{6.170a}$$

$$\frac{\mathrm{d}u}{\mathrm{d}\ell} = \epsilon u - 4\left[(n+8)u^2 + 6uv\right]\frac{1}{(1+r)^2}, \tag{6.170b}$$

$$\frac{\mathrm{d}v}{\mathrm{d}\ell} = \epsilon v - 4\left[12uv + 9v^2\right]\frac{1}{(1+r)^2}. \tag{6.170c}$$

Exercise 6.7 Obtain these flow equations.

We have proceeded to an appropriate redefinition of u and v in order to get rid of the K_d factor. These equations have the following four fixed points.

Gaussian: $u_{\text{G}}^* = v_{\text{G}}^* = 0$.

Heisenberg: $u_{\text{H}}^* = \epsilon/[4(n+8)]$, $v_{\text{H}}^* = 0$.

Ising: $u_{\text{I}}^* = 0$, $v_{\text{I}}^* = \epsilon/36$.

Cubic: $u_{\text{C}}^* = \epsilon/(12n)$, $v_{\text{C}}^* = [(n-4)\epsilon]/(36n)$.

The linearized flow equations in the proximity of the fixed points have the form

$$\frac{\mathrm{d}\delta u}{\mathrm{d}\ell} = \left\{\epsilon - 4\left[2(n+8)u^* + 6v^*\right]\right\}\delta u - 24u^*\,\delta v, \tag{6.171a}$$

$$\frac{\mathrm{d}\delta v}{\mathrm{d}\ell} = \left\{\epsilon - 4\left[12u^* + 18v^*\right]\right\}\delta u - 48v^*\,\delta v. \tag{6.171b}$$

The corresponding stability exponents are as follows.

Gaussian: $\lambda_{\text{G}}^u = \lambda_{\text{G}}^v = \epsilon$.

Heisenberg: $\lambda_{\text{H}}^u = -\epsilon$, $\lambda_{\text{H}}^v = (n-4)\epsilon/(n+8)$.

Ising: $\lambda_{\text{I}}^u = \epsilon/3$, $\lambda_{\text{I}}^v = -\epsilon$.

Cubic: $\lambda_{\text{C}}^{(1)} = -\epsilon$, $\lambda_{\text{C}}^{(2)} = (4-n)\epsilon/(3n)$.

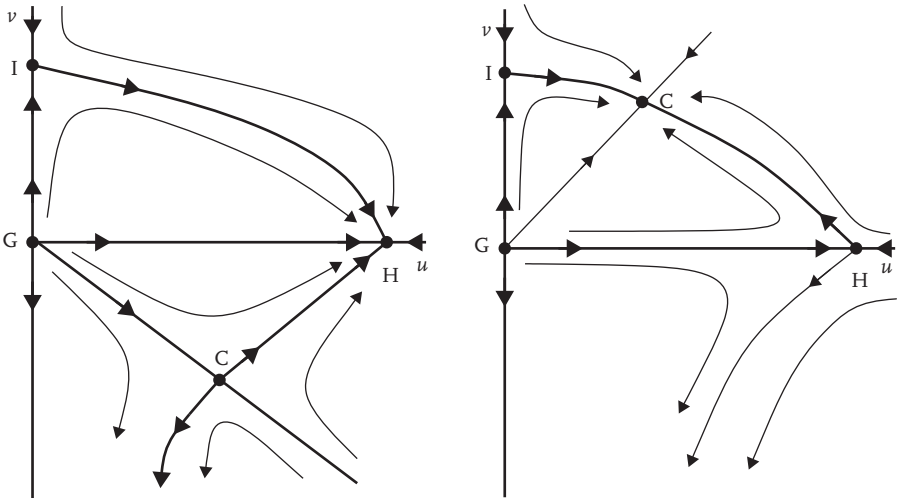


FIGURE 6.9. Renormalization flow in the (u, v) plane for the ϕ^4 model with cubic anisotropy. The following fixed points are shown: (G) Gaussian; (H) Heisenberg; (I) Ising; and (C) cubic. Left: $n < n_c$. Right: $n > n_c$.

One should note that the Heisenberg fixed point is unstable and the cubic fixed point is stable for $n > n_c = 4 + O(\epsilon)$, while the opposite is true for $n < n_c$. These equations imply the flow diagrams shown in figure 6.9—for all values of n, there is a fixed point that is locally stable. For $n < n_c$, the attraction basin of the H fixed point is the whole region above the G-C line. For $n > n_c$, the whole region $v > 0$ is the attraction basin of C, while only the line $v = 0$ is the attraction basin of H. There exist regions of the plane (u, v), however, that do not go to a fixed point—the entire region below the G-C line for $n < n_c$ and the region $v < 0$ for $n > n_c$. This phenomenon is known as **runaway**. In order to define the behavior of a system in which runaway appears, further analysis is necessary. In our case, it is possible to show that the transition becomes discontinuous. Thus a transition that appears continuous in mean-field theory turns out to be discontinuous because of fluctuations. This phenomenon is known as the Coleman-Weinberg mechanism [Cole73], and is thoroughly discussed in [Amit05, sec. 4-3].

By exploiting the flow eqs. (6.170), we can evaluate the exponent ν around the new cubic fixed point:

$$\nu_C = \frac{1}{2} + \frac{n-1}{6n}\epsilon + O\left(\epsilon^2\right). \tag{6.172}$$

6.10 Lower and Upper Critical Dimensions

One of the main lessons that we learned from the calculations we just reviewed is that it is useful to consider the dimensionality d of the system as a variable. We have obtained some nonclassical critical exponents for $d < 4$. For $d > 4$, the Ginzburg criterion guarantees us that the exponents are the same as the classical ones. The same result can be obtained by observing that if $\epsilon < 0$, the recurrence relations tell us that the stable fixed point is the

Gaussian one, where $u = 0$. Let us note that this does not imply that the exponents are those from the Gaussian model—in fact, the model is not even defined for $r < 0$. Let us, for instance, consider the behavior of the order parameter. Since one has

$$m \propto \left(\frac{-r}{u}\right)^{1/2}, \qquad r < 0, \tag{6.173}$$

according to Landau theory, and since $u \to 0$ under the renormalization flow, the singularity in the dependence on u changes the asymptotic behavior. The result is that the exponents are the classical ones, not the Gaussian ones, and especially that $\beta = 1/2$, instead of $(d-2)/4$, which would be what one would expect following the scaling laws.

The dimension d_c, below which exponents differ from the classical ones, is called the **upper critical dimension**. Since we may occasionally have to deal with problems that differ from the "usual" critical phenomena, it is useful to have at one's disposal a criterion to identify the upper critical dimension. Such a criterion was proposed by Toulouse [Toul75].

When the Kadanoff transformations admit of a nontrivial fixed point, the scaling laws apply, as we saw. They can be divided into two classes: those like $\alpha + 2\beta + \gamma = 2$, in which the dimensionality does not appear explicitly, and those like $2 - d\nu = \alpha$, in which it appears. The first are satisfied by the classical exponents; the second (generally) are not, since the classical exponents do not depend on dimensionality. The nonclassical exponents, on the other hand, satisfy the second group of scaling laws for all the values of d for which they are valid. At the upper critical dimension, the nontrivial fixed point is confused with the Gaussian (free of interaction), and the classical exponents therefore must satisfy the scaling laws: by setting $\alpha_c = 0$, $\nu_c = 1/2$, we obtain

$$d_c = \frac{2 - \alpha_c}{\nu_c} = 4. \tag{6.174}$$

For example, if a term proportional to ϕ^3 were present in the expansion of Landau free energy, one would obtain the classical exponents $\alpha = -1$, $\nu = 1/2$, and $d_c = 6$ would therefore follow. Since strictly speaking these terms are present in the van der Waals equation, we can ask ourselves why the Ising exponents that we obtained for $n = 1$ apply to the liquid–gas transition. The answer is that with a scalar order parameter it is always possible to eliminate these terms by redefining ϕ by means of a translation: $\phi \longrightarrow \phi + \phi_0$. When this transformation is forbidden because of the problem's symmetry, however, we are dealing with a new universality class and we obtain some new exponents, which must be calculated in an expansion in $6 - d$. This is the case of the percolation transition, which we discuss in the final chapter of this book. Analogously, in the case of the tricritical point (see the exercise at the end of sec. 5.15), where a symmetry-breaking transition changes from continuous to discontinuous, in the relevant fixed point the coefficient of the ϕ^4 theory vanishes and the ϕ^6 interaction becomes relevant. Here, the upper critical dimension is equal to 3, and therefore in real systems logarithmic corrections appear close to the tricritical point.

Exercise 6.8 Obtain explicitly the classical exponents for the ϕ^3 and the ϕ^6 cases and evaluate d_c from the Toulouse criterion.

Another remarkable value of dimensionality is the one where symmetry breaking occurs at zero temperature. This is called the **lower critical dimension**. For $n = 1$ (Ising), it is equal to 1, as can be seen by Landau's argument (see sec. 5.10.1), while for $n \geq 2$, where the symmetry is continuous, it is equal to 2. We can convince ourselves of this fact by considering a system in which the order parameter $\boldsymbol{\phi}$ has approximately uniform magnitude across the sample, but a varying direction. Then the excess free energy $\mathcal{F}[\boldsymbol{\phi}]$ of such a configuration is proportional to $\int \mathrm{d}\boldsymbol{r}\ |\nabla\boldsymbol{\phi}|^2$. By averaging over the Gaussian Boltzmann factor $\mathrm{e}^{-\mathcal{F}[\boldsymbol{\phi}]/k_\mathrm{B}T}$, we obtain that the correlation $\langle\boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}}\rangle$ is proportional to $k_\mathrm{B}T/k^2$. Evaluating the contribution to the local fluctuations due to these orientational fluctuations, we obtain

$$\left\langle |\delta\boldsymbol{\phi}|^2 \right\rangle = \int \mathrm{d}\boldsymbol{k}\ \langle\boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}}\rangle \propto \int \mathrm{d}\boldsymbol{k}\ \frac{k_\mathrm{B}T}{k^2} \propto k_\mathrm{B}T \int_{\pi/L}^{\pi/a_0} k^{d-3}\,\mathrm{d}k, \tag{6.175}$$

where a_0 is the lattice constant and L is the linear size of the system. For $d > 2$ this contribution remains finite as L goes to infinity. We can also use this approach to obtain the power-law behavior of the correlation function as a function of the distance $\boldsymbol{r}$ in the broken-symmetry phase:

$$\left\langle \delta\boldsymbol{\phi}(0) \cdot \delta\boldsymbol{\phi}(\boldsymbol{r}) \right\rangle = \int \mathrm{d}\boldsymbol{k}\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \langle\boldsymbol{\phi}_{\boldsymbol{k}} \cdot \boldsymbol{\phi}_{-\boldsymbol{k}}\rangle \sim |\boldsymbol{r}|^{2-d}\,. \tag{6.176}$$

For $d = 2$ we obtain instead a contribution to $\left\langle |\delta\boldsymbol{\phi}^2| \right\rangle$ proportional to $k_\mathrm{B}T \ln L$. Therefore, for $T > 0$, the orientational fluctuations destroy any possibility of order if the sample is large enough. Since the divergence is only logarithmic, however, this effect can only be seen in fairly large samples. This reasoning can be made rigorous [Merm66]. For $n \geq 2$, it is possible to expand the critical exponents into powers of $d - 2$.

6.11 Kosterlitz-Thouless Transition

For $n = 2$ and $d = 2$, even though there is no symmetry breaking, there is a peculiar phase transition at finite temperature, which is called the **Kosterlitz-Thouless transition**.

It is useful to parametrize the fluctuating field $\boldsymbol{\phi}$ of an $n = 2$ system as a complex number:

$$\phi_1(\boldsymbol{r}) + \mathrm{i}\phi_2(\boldsymbol{r}) = \phi_0\, \mathrm{e}^{\mathrm{i}\theta(\boldsymbol{r})}, \tag{6.177}$$

where $\phi_0 > 0$ is assumed to be essentially constant in what follows. Consider now a configuration in which $\theta(\boldsymbol{r})$ is slowly varying. Then the increment in free energy with respect to a uniform configuration is approximately given by

$$\mathcal{F}[\theta] = \frac{\kappa}{2} \int \mathrm{d}\boldsymbol{r}\ \left|\nabla\theta(\boldsymbol{r})\right|^2, \tag{6.178}$$

where κ is a positive constant. Going to the Fourier transform, we obtain

$$\mathcal{F}[\theta] = \frac{\kappa}{2}\, V \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2}\, k^2\, \theta_{\boldsymbol{k}}\theta_{-\boldsymbol{k}}, \tag{6.179}$$

where $V = L^2$ is the system's volume. We now momentarily neglect the fact that θ is constrained, say, between $-\pi$ and π. We wish to evaluate the correlation function

$$C(\boldsymbol{r}) = \langle \boldsymbol{\phi}(0) \cdot \boldsymbol{\phi}(\boldsymbol{r}) \rangle = \phi_0^2 \, \mathrm{Re} \left\langle \mathrm{e}^{\mathrm{i}(\theta(0) - \theta(\boldsymbol{r}))} \right\rangle, \tag{6.180}$$

where the average is taken with respect to the Gaussian distribution $\mathrm{e}^{-\mathcal{F}[\theta]/k_\mathrm{B}T}$. By using the Fourier antitransformation

$$\theta(\boldsymbol{r}) = \frac{1}{\sqrt{V}} \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2} \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \, \theta_{\boldsymbol{k}}, \tag{6.181}$$

and performing the Gaussian integration over the $\theta_{\boldsymbol{k}}$, we arrive at

$$C(\boldsymbol{r}) = \exp\left[-\frac{k_\mathrm{B}T}{\kappa} \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2} \, \frac{1 - \cos(\boldsymbol{k}\cdot\boldsymbol{r})}{k^2} \right]. \tag{6.182}$$

We perform the integral over $\boldsymbol{k}$ in polar coordinates, setting the cutoff at $|\boldsymbol{k}| = \pi/a_0$, where a_0 is the lattice constant, and taking advantage of the relation

$$\int_0^{2\pi} \mathrm{d}\vartheta \, \cos(x \cos\vartheta) = 2\pi \, J_0(x), \tag{6.183}$$

where $J_0(x)$ is the zeroth-order Bessel function. We thus have

$$C(\boldsymbol{r}) = \exp\left[-\frac{k_\mathrm{B}T}{2\pi\,\kappa} \int_0^{\pi/a_0} \mathrm{d}k \, \frac{1 - J_0(kr)}{k} \right]. \tag{6.184}$$

Changing the integration variable to $x = kr$ we obtain the integral

$$\int_0^{\pi r/a_0} \mathrm{d}x \, \frac{1 - J_0(x)}{x}.$$

For large values of r/a_0 the integral is dominated by the large x values, where $J_0(x)$ is small. We thus have

$$C(\boldsymbol{r}) \simeq \exp\left[-\frac{k_\mathrm{B}T}{2\pi\kappa} \ln\left(\frac{\pi r}{a_0}\right) \right] = \left(\frac{\pi r}{a_0}\right)^{-k_\mathrm{B}T/(2\pi\kappa)} \propto r^{-\eta(T)}. \tag{6.185}$$

We have thus obtained the surprising result that, at low temperatures, although there is no spontaneous symmetry breaking, the order parameter correlation function decays as a power law, with a temperature-dependent exponent η. One then speaks of **quasi long-range order**. On the other hand, we expect that at high enough temperatures the correlation function decays exponentially, with a well-defined coherence length ξ. Therefore, a phase transition should separate these two regimes.

Which mechanism lies at the basis of this transition? In the above argument, we have neglected the fact that the state of the system does not change under the shift of θ by an integer multiple of 2π. It is therefore possible for the system to sustain configurations such that, by integrating $\nabla\theta$ over a closed loop, one obtains a nontrivial integer value of 2π. Two configurations such that $\oint \nabla\theta \cdot \mathrm{d}\boldsymbol{r} = n$ on any path circling the central dot, with $n = 1$ (a **vortex**, left) and $n = -1$ (an **antivortex**, right), are shown in fig. 6.10. These

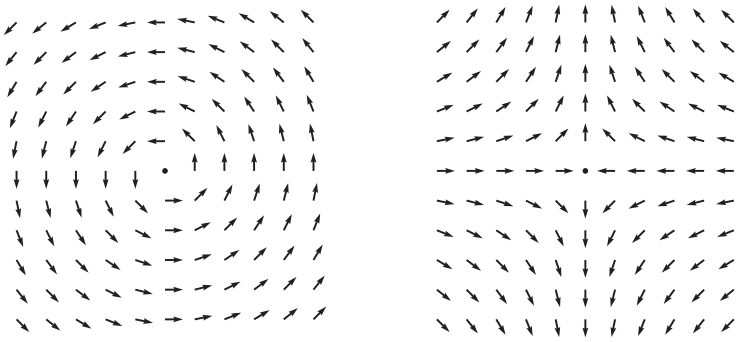


FIGURE 6.10. Example of a vortex (left) and an antivortex (right) configuration.

configurations necessarily imply that there is a small region at their center where the phase θ is not well defined, which is called the **singularity**, and the corresponding value of n is called its **strength**.

Let us evaluate the free energy cost of such configurations. It is clear that $|\nabla\theta| = (2\pi\,|n|)/r \propto 1/r$, where r is the distance from the singularity (marked by a black dot in the figures). Thus, setting the singularity at the origin and integrating in polar coordinates, we obtain

$$\mathcal{F} = \frac{\kappa}{2}\int \mathrm{d}\boldsymbol{r}\;|\nabla\theta|^2 = \pi\kappa n^2 \int_{a_0}^{L} \frac{\mathrm{d}r}{r} = \pi\kappa n^2 \ln\frac{L}{a_0}, \tag{6.186}$$

where L is of the order of the linear size of the system and a_0 is the lattice constant. On the other hand, the entropy of an isolated singularity is given by $k_{\mathrm{B}}\ln(L/a_0)^2$. It follows that there is a finite value of the temperature $T_{\mathrm{c}} = \pi\kappa n^2/(2k_{\mathrm{B}})$ above which the entropy gain in producing vortices overcomes the corresponding energy loss. For temperatures above this one, the system will be covered by a plasma of vortices and antivortices, and we expect the correlation function of θ to decay over a distance of the order of the inverse square root of the density of singularities.

However, the physics of this transition is even subtler. Indeed, while the energy cost of an isolated singularity diverges with the system size, the energy of a matched pair of singularities with opposite values of n (for example, a vortex and an antivortex) is finite and grows logarithmically with the distance between them. Indeed, the minimum free-energy configuration of $\theta(\boldsymbol{r})$ satisfies the condition $\delta\mathcal{F}/\delta\theta(\boldsymbol{r}) = 0$, which corresponds to

$$\nabla^2\theta(\boldsymbol{r}) = 0. \tag{6.187}$$

On the other hand, the configuration must satisfy the condition

$$\oint_{\mathcal{C}} \nabla\theta\cdot\mathrm{d}\boldsymbol{r} = 2\pi n, \tag{6.188}$$

where the contour $\mathcal{C}$ encircles a singularity of strength n. Obviously, this condition cannot be satisfied if θ is a single-valued function. Let us therefore denote the distortion field $\nabla\theta$

by $\boldsymbol{\psi}$. By Stokes's theorem we have

$$2\pi n = \oint_{\mathcal{C}} \boldsymbol{\psi} \cdot \mathrm{d}\boldsymbol{r} = \int_{\mathcal{S}} (\nabla \times \boldsymbol{\psi}) \cdot \boldsymbol{e}_z \, \mathrm{d}^2\boldsymbol{r}, \tag{6.189}$$

where $\mathcal{S}$ is the portion of the (x, y) plane bounded by the closed contour $\mathcal{C}$. Here $\boldsymbol{e}_z$ is the versor of the z axis and $\nabla \times \boldsymbol{\psi}$ is a vector parallel to it. Therefore, for a singularity of strength n placed at the origin, we have

$$\nabla \times \boldsymbol{\psi}(\boldsymbol{r}) = 2\pi \; n\, \delta(\boldsymbol{r})\, \boldsymbol{e}_z. \tag{6.190}$$

Let us define the Green's function $G(\boldsymbol{r})$ as the solution of the Poisson equation

$$-\nabla^2 G(\boldsymbol{r}) = \delta(\boldsymbol{r}). \tag{6.191}$$

Then the optimal configuration is given by

$$\boldsymbol{\psi}_n(\boldsymbol{r}) = 2\pi n \, \nabla \times G(\boldsymbol{r})\, \boldsymbol{e}_z, \tag{6.192}$$

where we have used the identity

$$\nabla \times (\nabla \times \boldsymbol{A}) = \nabla\,(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}, \tag{6.193}$$

in which the first term vanishes when $\boldsymbol{A} = A(x, y)\, \boldsymbol{e}_z$. Since eq. (6.187) is linear, in the presence of several singularities α of strength n_α placed at $\boldsymbol{r}_\alpha$ we have the optimal configuration

$$\boldsymbol{\psi}^*(\boldsymbol{r}) = \sum_\alpha \boldsymbol{\psi}_{n_\alpha}(\boldsymbol{r} - \boldsymbol{r}_\alpha) = 2\pi \; \nabla \times \sum_\alpha n_\alpha \; G(\boldsymbol{r} - \boldsymbol{r}_\alpha)\, \boldsymbol{e}_z. \tag{6.194}$$

The corresponding free energy cost is given by

$$\begin{aligned}
\mathcal{F}[\boldsymbol{\psi}^*] &= \frac{\kappa}{2} \int \mathrm{d}\boldsymbol{r} \; \left|\boldsymbol{\psi}^*\right|^2 \\
&= \frac{\kappa}{2} (2\pi)^2 \sum_{\alpha\beta} n_\alpha n_\beta \int \mathrm{d}\boldsymbol{r} \; (\nabla \times G(\boldsymbol{r} - \boldsymbol{r}_\alpha)\boldsymbol{e}_z) \cdot \left(\nabla \times G(\boldsymbol{r} - \boldsymbol{r}_\beta)\boldsymbol{e}_z\right) \\
&= -\frac{\kappa}{2} (2\pi)^2 \sum_{\alpha\beta} n_\alpha n_\beta \int \mathrm{d}\boldsymbol{r} \; G(\boldsymbol{r} - \boldsymbol{r}_\alpha)\, \boldsymbol{e}_z \left[\nabla \times \left(\nabla \times G(\boldsymbol{r} - \boldsymbol{r}_\beta)\boldsymbol{e}_z\right)\right] \\
&= \frac{\kappa}{2} (2\pi)^2 \sum_{\alpha\beta} n_\alpha n_\beta \int \mathrm{d}\boldsymbol{r} \; G(\boldsymbol{r} - \boldsymbol{r}_\alpha) \delta(\boldsymbol{r} - \boldsymbol{r}_\beta) \\
&= \frac{\kappa}{2} (2\pi)^2 \sum_{\alpha\beta} n_\alpha n_\beta \; G(\boldsymbol{r}_\beta - \boldsymbol{r}_\alpha).
\end{aligned} \tag{6.195}$$

Now for $|\boldsymbol{r}| \gg a_0$ we have

$$G(\boldsymbol{r}) = \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^2} \, \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}}{k^2} \simeq \frac{1}{2\pi} \ln \frac{L}{a_0} - \frac{1}{2\pi} \ln \frac{|\boldsymbol{r}|}{a_0} + \text{const.} \tag{6.196}$$

The first term yields a contribution proportional to $(\sum_\alpha n_\alpha)^2$ that diverges for $L \to \infty$. Thus cases where $\sum_\alpha n_\alpha \neq 0$ are forbidden. Moreover, the terms with $\alpha = \beta$ yield a contribution given by $\sum_\alpha n_\alpha^2$ times the energy associated with the core of the singularity, which is finite. We are thus left with

$$\mathcal{F}[\boldsymbol{\psi}^*] \simeq -\kappa\pi \sum_{\alpha\neq\beta} n_\alpha n_\beta \ln \left| \frac{\boldsymbol{r}_\alpha - \boldsymbol{r}_\beta}{a_0} \right| + \text{const.} \tag{6.197}$$

When n_α and n_β have the same sign, the interaction is repulsive, otherwise it is attractive. In both cases it grows in magnitude with the logarithm of the distance. (It is worthwhile to remark that this interaction looks like that of Coulomb charges of magnitude $2\pi n$ interacting in two dimensions in a medium of permittivity κ.) Therefore, for small values of T, we expect the system to contain a certain number of closely bound vortex-antivortex pairs. The dipole interaction between these pairs falls off fast enough to be irrelevant. The presence of these pairs locally disrupts the configuration of the order parameter $\boldsymbol{\phi}$, leading to a decrease in the effective elasticity κ at long distances. As the temperature increases, more bound vortex-antivortex pairs appear and, moreover, become less closely bound. At a certain temperature, the expected distance between a vortex and its companion will be of the same order as that between a pair and its closest neighboring pair. When this happens, the pairs become unbound, and we are left with the plasma phase we mentioned earlier.

Exercise 6.9 You may convince yourself of the analogy between vortices and charges by showing that, in $2d$, the circulation of a vector field $\boldsymbol{V}(\boldsymbol{r})$ over a closed circuit $\mathcal{C}$ is equal to the flux of a suitably rotated vector field $\boldsymbol{W}(\boldsymbol{r})$ across the same circuit $\mathcal{C}$. Using this analogy, obtain the above results by using Coulomb's law and the divergence theorem.

This physical picture, due to V. Berezinskii [Bere71], J. M. Kosterlitz, and D. J. Thouless [Kost73], can be made more explicit by a renormalization group treatment, which is described in [Jose77] or [Nels83]. (A detailed pedagogical treatment can be found in [Chai95, chap. 9].) While I am not going to enter into the technical details, it is possible to explain this approach in the following way.

At low temperatures, where the number of singularities is small, one can expand the partition function in powers of the fugacity z of the singularities, interacting via a medium of effective stiffness $K = \kappa/k_B T$. The first contribution will be proportional to z^2, corresponding to a vortex-antivortex pair, and will be of the form

$$z^2 \int \frac{\mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2}{a_0^4}\, \mathrm{e}^{-E(\boldsymbol{r}_1,\boldsymbol{r}_2)/k_B T},$$

where, according to eq. (6.197), the energy of such a pair is given by

$$E(\boldsymbol{r}_1, \boldsymbol{r}_2) = 2\pi\,\kappa \ln \left| \frac{\boldsymbol{r}_1 - \boldsymbol{r}_2}{a_0} \right| + \text{const.}, \tag{6.198}$$

where the constant represents the core energy. We can get rid of it by suitably rescaling the fugacity z.

One then proceeds at renormalization by grouping together lattice points, that is, by increasing the lattice constant a_0 by a factor e^{ℓ}, and then by rescaling all lengths. From eq. (6.198) we see that the energy change of a vortex pair, as a_0 is rescaled by $e^{\delta\ell}$, is given by $2\,\delta E = 2\pi\kappa\,\delta\ell$. This corresponds to a change $\delta \ln z = -\delta E/k_B T$ in the fugacity. Taking into account the density rescaling that yields a contribution $\delta \ln z = 2\,\delta\ell$, we obtain

$$\frac{\delta z}{z} = \left(2\,\delta\ell - \frac{\delta E}{k_B T}\right) = (2 - \pi K)\;\delta\ell. \tag{6.199}$$

Let us denote by x the expression $2 - \pi K$. The effective stiffness will be decreased by taking into account the vortex-antivortex pairs that are erased by the renormalization. Thus, to lowest order, we have a change in x that is proportional to z^2, since the linear term is ruled out by the symmetry $z \longrightarrow -z$ that ensures that vortices always appear paired with antivortices. Therefore, summing up, we obtain renormalization flow equations to lowest order in x and z of the form

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}\ell} &= Az^2 + \cdots, \\ \frac{\mathrm{d}z}{\mathrm{d}\ell} &= xz + \cdots, \end{aligned} \tag{6.200}$$

where A is a positive constant. No other terms are allowed in the equations to lowest order in x and z. These equations allow for a line of fixed points at $z = 0$, which is stable for $x < 0$, that is, for $K^{-1} < 1/K^* = \pi/2$, and unstable otherwise. Thus we can envisage the renormalization flow in the space spanned by K^{-1} (which is proportional to T) and z. If z is small enough, or K large enough, as we keep on renormalizing, z will decrease to a point where there are essentially no more singularities in the system. At this point the system is described by a quadratic Hamiltonian of the form $\int \mathrm{d}\boldsymbol{r}\, K\,|\nabla\theta|^2\,/2$, and the stiffness does not change anymore. We have therefore a line of fixed points, each of them with correlations decaying with K-dependent exponents, that corresponds to a quasi long-range order. If we start with a larger value of z (or a smaller value of K), then, before reaching this regime, the stiffness will have decreased to a point such that the distance between singularities in a bound vertex-antivertex pair is of the order of the distance between pairs, and the system will switch to the plasma phase. In this phase, we expect z to increase and K to decrease even faster upon renormalization.

We thus have a critical line in the (K^{-1}, z) plane that separates a quasi long-range ordered phase from the plasma phase. This line ends at point $(1/K^*, 0)$. For $K < K^*$, the line of fixed points $(1/K, 0)$ becomes unstable. Interestingly, this means that the effective stiffness jumps from a finite and universal value K^* to zero at the transition, while there is no detectable specific heat singularity. A qualitative sketch of the renormalization flow is shown in fig. 6.11. The theory has been tested with success in superfluid ^{4}He films [Bish78] and has been extended to describe two-dimensional melting by J. M. Kosterlitz, D. J. Thouless, B. Halperin, D. R. Nelson, and A. P. Young (see [Nels83]). Surprisingly, melting in two-dimensional crystals takes place in two steps: first, the system loses quasi long-range positional order, but keeps quasi long-range order in the orientation of the six bonds connecting each atom to its nearest neighbors. This is called the **hexatic phase**, since the

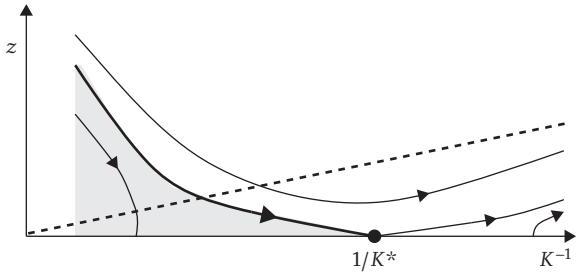


FIGURE 6.11. Renormalization flow for the Kosterlitz-Thouless transition (schematic). The thick line is the trajectory that runs into the fixed point $(1/K^*, 0)$, and separates the quasi long-range order region below (in gray) from the plasma phase above. The line of fixed points that describes quasi long-range order corresponds to $z = 0$ and $0 < K^{-1} \leq 1/K^*$. The dashed line is the trajectory that the Hamiltonian follows as T varies, and acts as a line of starting points for the renormalization procedure. The critical temperature corresponds to the crossing of the dashed and the thick lines.

orientational order is sixfold, that is, maps onto itself by a rotation of 60°. In this first transition, the role of vortices is played by lattice dislocations. Then a second transition of the Kosterlitz-Thouless type destroys the remaining order, leading to an isotropic fluid phase by the unbinding of disclination pairs, where a disclination is the analog of a vortex with a charge multiple of 1/6.

Recommended Reading

Various texts contain reviews of the renormalization group at differing levels of detail. A classic text is D. J. Amit and V. Martín-Mayor, *Field Theory, the Renormalization Group, and Critical Phenomena*, 3rd ed., Singapore: World Scientific, 2005. A focused and concise discussion can be found in J. Cardy, *Scaling and Renormalization in Statistical Physics*, Cambridge, UK: Cambridge University Press, 1996. One can also profitably read J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. E. J. Newman, *The Theory of Critical Phenomena*, Oxford, UK: Clarendon Press, 1992; chapter 6 of M. Plischke and B. Birgersen, *Equilibrium Statistical Physics*, 3rd ed., Singapore: World Scientific, 2006; and N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*, Boca Raton, FL: CRC Press, 1992. A concise but very accessible introduction to the field theoretical approach to critical phenomena is É. Brézin, *Introduction to Statistical Field Theory*, Cambridge, UK: Cambridge University Press, 2010. The text by H. Kleinert and V. Schulte-Frohlinde, *Critical Properties of ϕ^4 Theories*, Singapore: World Scientific, 2001, contains an extensive analysis of the field theory approach, with exponent evaluations at high order in perturbation theory.

7 Classical Fluids

We shall treat the armor's steel as an ideal fluid.
—Lavrentiev and Shabat

The first application of statistical mechanics was, historically speaking, the study of classical fluids (see, for example, Boltzmann's *Gastheorie* [Bolt95]). In recent years, it has become the domain of a specialized and quite sophisticated community of researchers. In particular, there has been a lot of interest in the study of *complex fluids*, such as those composed of mixtures of several chemical species, anisotropic (liquid crystals) or nonhomogeneous (interfaces and microemulsions). Water is also considered a complex fluid, because of its special physical characteristics, which are determined by the hydrogen bond. More generally, the field of *soft matter* has been developing vigorously in the last decades: under this label, one regroups a great variety of physical systems that includes colloids, polymers, foams, gels, and a number of biological materials. They all consist of units of intermediate size between the atomic and the macroscopic ones. As a consequence, a great number of their properties are shaped by thermal fluctuations, and their response to external perturbations is typically large, nonlinear, and slow. Unfortunately, it is beyond the scope of this book to account for these developments.

My only goal is to provide an introduction to the fundamental methods of the theory of simple fluids, namely, those composed of a single molecular species. In fact, by combining analytical and numerical methods, and on the basis of a microscopic model, a fairly satisfying description of the phase diagram of simple systems like argon has been reached. In the present chapter I sketch the basic ideas underlying this theory.

7.1 The Partition Function of a Classical Fluid

Let us consider a model of a simple fluid, represented by a collection of N particles of mass m, placed in a container of volume V, and interacting by means of a pair potential $u(r)$ that depends only on the magnitude $r = |\boldsymbol{r}|$ of the distance $\boldsymbol{r}$ between particles. This model is already to a significant degree an idealization of a simple fluid. In effect:

- the internal degrees of freedom of the particles are disregarded,
- direct interactions between three or more bodies are excluded—such interactions are instead rather frequent. An important case is that of the ubiquitous van der Waals interactions that originate in dipole-dipole or multipole-multipole interactions, and contain interaction terms involving two, three, or more bodies.
- One neglects the dependence of the interaction on the orientation of the molecules, an effect that can lead to the formation of an anisotropic liquid.

This model, however, is already fairly useful for describing the behavior of fluids such as rare gases (excluding helium, whose properties are influenced by quantum effects) and some molecular gases such as CO_2.

We describe the interaction between the system and the environment by introducing a single-particle potential $v(\boldsymbol{r})$ acting on each particle, where $\boldsymbol{r} \in \mathbb{R}^3$ denotes the generic position of the particle. Then the Hamiltonian of the system has the expression

$$\mathcal{H}(x) = \sum_{i=1}^{N} \frac{p_i^2}{2m} + U\left(\boldsymbol{r}^N\right) + \sum_{i=1}^{N} v(\boldsymbol{r}_i), \tag{7.1}$$

where $x = (\boldsymbol{p}^N, \boldsymbol{r}^N)$ is the phase-space point, $\boldsymbol{p}^N = (\boldsymbol{p}_i)$ is the collection of the single-particle momenta $\boldsymbol{p}_i$, $\boldsymbol{r}^N = (\boldsymbol{r}_i)$ is the collection of the single-particle position vectors $\boldsymbol{r}_i$, and the potential energy $U\left(\boldsymbol{r}^N\right)$ is given by

$$U\left(\boldsymbol{r}^N\right) = \sum_{(i,j)} u\left(\left|\boldsymbol{r}_i - \boldsymbol{r}_j\right|\right), \tag{7.2}$$

where the sum runs over all distinct particle pairs (i,j). In this chapter, we consider only three-dimensional systems, unless explicitly stated otherwise.

The thermodynamic properties of this system can be obtained starting from the canonical partition function Z_N, which, taking into account the *Gibbs correction factor* and the *volume of the phase space element* (see section 3.8), is expressed as follows:

$$Z_N = \frac{1}{N!} \int \prod_{i=1}^{N} \left(\frac{\mathrm{d}\boldsymbol{p}_i \, \mathrm{d}\boldsymbol{r}_i}{h^3} \right) \exp\left[-\frac{\mathcal{H}(x)}{k_\mathrm{B} T} \right]. \tag{7.3}$$

The Gaussian integral over the momenta can be evaluated directly, so that one obtains

$$Z_N = \frac{1}{N!} \int \prod_{i=1}^{N} \left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_\mathrm{B}^3} \right) \exp\left\{ -\frac{1}{k_\mathrm{B} T} \left[U\left(\boldsymbol{r}^N\right) + \sum_{i=1}^{N} v(\boldsymbol{r}_i) \right] \right\} = \frac{1}{N!} \lambda_\mathrm{B}^{-3N} Q_N[v], \tag{7.4}$$

where λ_B is the thermal de Broglie wavelength

$$\lambda_\mathrm{B} = \sqrt{\frac{h^2}{2\pi m k_\mathrm{B} T}}, \tag{7.5}$$

and we have defined the **configuration integral**

$$Q_N[v] = \int \prod_{i=1}^{N} \mathrm{d}\boldsymbol{r}_i \, \exp\left\{ -\frac{1}{k_\mathrm{B} T} \left[U\left(\boldsymbol{r}^N\right) + \sum_{i=1}^{N} v(\boldsymbol{r}_i) \right] \right\}. \tag{7.6}$$

The notation emphasizes that Q_N is a functional of the one-particle potential $v(\boldsymbol{x})$. We shall see later that this observation has important consequences, since it leads to the *density functional theory*.

It is often useful to deal with the system in the grand canonical ensemble by introducing the chemical potential μ. The grand canonical partition function Z_{GC} is then given by

$$Z_{\mathrm{GC}} = \sum_{N=0}^{\infty} \mathrm{e}^{\mu N/k_{\mathrm{B}}T} Z_N = \sum_{N=0}^{\infty} \frac{1}{N!} \left(\frac{z}{\lambda_{\mathrm{B}}^3} \right)^N Q_N[v], \tag{7.7}$$

where $z = \mathrm{e}^{\mu/k_{\mathrm{B}}T}$ is the fugacity.

Now, let us consider in particular the case of noninteracting particles, for which $U\left(\boldsymbol{r}^N\right) = 0$. We then have

$$Z_{\mathrm{GC}} = \sum_{N=0}^{\infty} \frac{z^N}{N!} \left(Z_1[v]\right)^N = \exp\left(z\, Z_1[v]\right), \tag{7.8}$$

where $Z_1[v]$ is the canonical partition function of a single particle, expressed as a functional of the external potential v:

$$Z_1[v] = \int \frac{\mathrm{d}\boldsymbol{r}}{\lambda_{\mathrm{B}}^3} \mathrm{e}^{-v(\boldsymbol{r})/k_{\mathrm{B}}T}. \tag{7.9}$$

Therefore, if $v(\boldsymbol{r})$ is a potential confining the system in a box $\mathcal{B}$ of volume V, that is, if

$$v(\boldsymbol{r}) = v_{\mathcal{B}}(\boldsymbol{r}) = \begin{cases} 0, & \text{if } \boldsymbol{r} \in \mathcal{B}, \\ +\infty, & \text{otherwise,} \end{cases} \tag{7.10}$$

one has $Z_1 = V/\lambda_{\mathrm{B}}^3$. Therefore $Z_{\mathrm{GC}} = \exp(zV/\lambda_{\mathrm{B}}^3)$, which allows one to obtain all the usual properties of ideal gases, as we saw in chapter 4.

In order to proceed, we need to specify some properties of the pair interaction potential $u(r)$. In the systems we are considering, it exhibits a **hard core repulsion**—in other words, a strong repulsion over short distances—and an **attractive tail**, which rapidly decays over distances that are slightly larger than the particle radius. The hard core repulsion is due the fact that, as soon as the electron clouds of different molecules overlap, the Pauli principle requires some of the electrons to go to highly excited states. Thus the energy of the interacting pair becomes large very rapidly as the molecules get close. The attractive tail is due to several mechanisms, but the most common is the interaction between permanent or induced electric dipoles and multipoles. It is interesting that the interaction potential is basically determined by only one scale of length, namely, the radius of the hard-core repulsion. This allows one to represent it fairly easily. One expression that has been particularly successful is the 12-6 potential, also called the **Lennard-Jones potential**, in which only two parameters appear: the energy scale ϵ_0 and the length scale r_0. It is defined by

$$u(r) = 4\epsilon_0 \left[\left(\frac{r_0}{r} \right)^{12} - \left(\frac{r_0}{r} \right)^{6} \right]. \tag{7.11}$$

It is easy to see that ϵ_0 represents the greatest depth of the attractive potential, while r_0 satisfies $u(r_0) = 0$. This potential exhibits a very fast increase for $r < r_0$ and a slower, but still

quite fast, decrease of the attractive tail for $r > r_0$, such that one has, for example, $u(2r_0) = -0.06\,\epsilon_0$. More particularly, one has $\lim_{r\to\infty} r^3 u(r) = 0$. This relation does not hold for the Coulomb or dipolar potentials, which are therefore considered *long range*. The minimum of the potential is reached at $r = 2^{1/6} r_0$.

Since the Lennard-Jones potential contains only two parameters, like the van der Waals equation of state, the equation of state of a Lennard-Jones fluid assumes a universal form in terms of the reduced physical parameters—in other words, when expressed as a function of critical temperature, pressure, and density. The validity of the Lennard-Jones potential can therefore be evaluated by comparing the equations of state for different fluids, and especially the values for their reduced critical temperature and density. If the Lennard-Jones potential were exact, the reduced critical temperature and density would be the same for all systems. It turns out that the agreement is not very satisfactory, in particular for lighter gases like H_2 or He, where quantum effects come into play. However, due to its simplicity, the Lennard-Jones potential is a good starting point to explore the thermodynamical properties of fluids.

7.2 Reduced Densities

We have seen, however, that one of the main advantages of statistical mechanics is that it allows us, in principle, to provide a detailed description of the behavior of thermodynamic systems that goes beyond the mere equation of state. It is possible, in particular, to evaluate the averages of a number of quasi-thermodynamic observables—that is, of extensive observables, which cannot be controlled from the exterior, and cannot therefore be counted as thermodynamic variables. In the case of fluids, a very broad class of these quantities can be expressed in terms of **reduced densities**—in other words, of the correlation functions of the local particle density. In this section, we define the reduced densities and see how they are connected to the thermodynamic properties of the fluid, and how in certain cases they can be measured by means of scattering experiments. I refer the reader to the following sections to see how these quantities can be evaluated by means of a systematic power-series expansion of the densities (**virial expansion**). This expansion can be obtained by treating the interaction potential as a perturbation and by suitably resumming the resulting series. In a more sophisticated approach, one considers the properties of a fluid of hard spheres as known, and expands only on the (much weaker) attractive part of the potential.

Define the local fluctuating density $\hat{\rho}_N(\boldsymbol{x})$ by

$$\hat{\rho}_N(\boldsymbol{x}) = \sum_{i=1}^{N} \delta(\boldsymbol{x} - \boldsymbol{r}_i). \tag{7.12}$$

Then the single-particle reduced density is simply the local average of the particle density:

$$\rho^{(1)}(\boldsymbol{x}) = \left\langle \hat{\rho}_N(\boldsymbol{x}) \right\rangle. \tag{7.13}$$

We shall often denote it simply by $\rho(\boldsymbol{x})$. Since all the particles are identical, $\rho^{(1)}(\boldsymbol{x})$ can be evaluated by calculating the relative weight of the configurations in which a single particle

is found in $\boldsymbol{x}$:

$$\rho^{(1)}(\boldsymbol{x}) = \frac{N}{Q_N[v]} \int \mathrm{d}\boldsymbol{r}^N \, \exp\left\{ -\frac{1}{k_\mathrm{B}T} \left[U\left(\boldsymbol{r}^N\right) + \sum_{i=1}^{N} v(\boldsymbol{r}_i) \right] \right\} \delta\left(\boldsymbol{x} - \boldsymbol{r}_1\right), \tag{7.14}$$

where we have set $\mathrm{d}\boldsymbol{r}^N = \prod_{i=1}^{N} \mathrm{d}\boldsymbol{r}_i$ for brevity. Obviously, if $v(\boldsymbol{r})$ is equal to the box potential $v_\mathcal{B}(\boldsymbol{r})$, then $\rho^{(1)}(\boldsymbol{x}) = N/V = \rho$, at least not too close to the walls.

We can express the n-particle reduced density in an analogous way:

$$\rho^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \left\langle \prod_{\alpha=1}^{n} \hat{\rho}_N(\boldsymbol{x}_\alpha) \right\rangle = \left\langle \sum_{i_1,\ldots,i_n} \prod_{\alpha=1}^{n} \delta\left(\boldsymbol{x}_\alpha - \boldsymbol{r}_{i_\alpha}\right) \right\rangle. \tag{7.15}$$

This yields

$$\begin{aligned} \rho^{(n)}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = {} & \frac{N!}{(N-n)!} \frac{1}{Q_N[v]} \\ & \times \int \mathrm{d}\boldsymbol{r}^N \, \exp\left\{ -\frac{1}{k_\mathrm{B}T} \left[U\left(\boldsymbol{r}^N\right) + \sum_{i=1}^{N} v(\boldsymbol{r}_i) \right] \right\} \prod_{\alpha=1}^{n} \delta\left(\boldsymbol{x}_\alpha - \boldsymbol{r}_\alpha\right). \end{aligned} \tag{7.16}$$

Exercise 7.1 Consider the Gibbs entropy associated with the one-particle reduced density $\rho^{(1)}$, defined by

$$\mathcal{S}(\rho^{(1)}) = -k_\mathrm{B} \int \mathrm{d}x \, \rho^{(1)}(x) \ln \rho^{(1)}(x),$$

where $x = (\boldsymbol{r}, \boldsymbol{p})$ denotes the single-particle state. Here $\rho^{(1)}$ is defined by

$$\rho^{(1)}(x) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \delta(x - x_i) \right\rangle,$$

where the average is taken with respect to the canonical distribution. Show that

$$\mathcal{S}(\rho^{(1)}) \geq \frac{S}{N},$$

where S is the thermodynamical entropy of the system.

7.2.1 Pair Distribution

In addition to the single-particle density $\rho^{(1)}(\boldsymbol{x}) \simeq N/V = \rho$, the reduced density that is most commonly of interest is the **pair density** $\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2)$, defined as the joint probability density of finding one particle in $\boldsymbol{x}_1$ and a different one in $\boldsymbol{x}_2$:

$$\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \left\langle \hat{\rho}_N(\boldsymbol{x}_1) \hat{\rho}(\boldsymbol{x}_2) \right\rangle = \left\langle \sum_{i \neq j} \delta(\boldsymbol{x}_1 - \boldsymbol{r}_i) \delta(\boldsymbol{x}_2 - \boldsymbol{r}_j) \right\rangle. \tag{7.17}$$

For $|\boldsymbol{x}_1 - \boldsymbol{x}_2| \to \infty$ it tends to the value $N(N-1)/V^2 \simeq \rho^2$.

We will be mostly interested in the correlation of the density fluctuations. We thus define the **correlation function** $G(\boldsymbol{x}_1, \boldsymbol{x}_2)$ by

$$G(\boldsymbol{x}_1, \boldsymbol{x}_2) = \left\langle [\hat{\rho}_N(\boldsymbol{x}_1) - \rho]\, [\hat{\rho}_N(\boldsymbol{x}_2) - \rho] \right\rangle. \tag{7.18}$$

Using the definition (7.12) of $\hat{\rho}_N(\boldsymbol{x})$ and isolating the $i = j$ term in the sum appearing in eq. (7.18), we obtain

$$G(\boldsymbol{x}_1, \boldsymbol{x}_2) = \rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) + \rho\, \delta(\boldsymbol{x}_1 - \boldsymbol{x}_2). \tag{7.19}$$

In a homogeneous fluid, $G(\boldsymbol{x}_1, \boldsymbol{x}_2)$ depends only on the distance $r = |\boldsymbol{x}_1 - \boldsymbol{x}_2|$ between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. We thus obtain

$$G(\boldsymbol{x}_1, \boldsymbol{x}_2) = \rho\, \delta(\boldsymbol{x}_1 - \boldsymbol{x}_2) + \rho^2\, [g(|\boldsymbol{x}_1 - \boldsymbol{x}_2|) - 1], \tag{7.20}$$

where $g(r)$ is called the **pair distribution**, and is defined, for $r = |\boldsymbol{x}_1 - \boldsymbol{x}_2| > 0$, by

$$\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \rho^2\, g(r). \tag{7.21}$$

In a fluid, the influence of a molecule placed at $\boldsymbol{x}_1$ on one placed at $\boldsymbol{x}_2$ is rapidly lost, as $r = |\boldsymbol{x}_1 - \boldsymbol{x}_2|$ is larger than a few molecular diameters. We thus have $\lim_{r\to\infty} g(r) = 1$. On the other hand, hard-core repulsion prevents two (or more) particles from coming closer than a molecular diameter: we have therefore $g(r) = 0$ for r smaller than a molecular diameter.

In a homogeneous and isotropic system in which only pair interactions are present, the average interaction energy can be expressed in terms of the $g(r)$:

$$\langle U \rangle = \frac{V\rho^2}{2} \int_0^\infty \mathrm{d}r\, 4\pi r^2\, u(r)\, g(r). \tag{7.22}$$

One has in fact

$$\begin{aligned}
\langle U \rangle &= \left\langle \frac{1}{2} \sum_{i \neq j} u\left(|\boldsymbol{r}_i - \boldsymbol{r}_j|\right) \right\rangle \\
&= \frac{N(N-1)}{2} \int \mathrm{d}\boldsymbol{x}_1\, \mathrm{d}\boldsymbol{x}_2\ \langle u(|\boldsymbol{r}_1 - \boldsymbol{r}_2|)\, \delta(\boldsymbol{x}_1 - \boldsymbol{r}_1)\, \delta(\boldsymbol{x}_2 - \boldsymbol{r}_2) \rangle \\
&= \frac{N(N-1)}{2} \int \mathrm{d}\boldsymbol{x}_1\, \mathrm{d}\boldsymbol{x}_2\, u(|\boldsymbol{x}_1 - \boldsymbol{x}_2|)\ \langle \delta(\boldsymbol{x}_1 - \boldsymbol{r}_1)\, \delta(\boldsymbol{x}_2 - \boldsymbol{r}_2) \rangle \\
&= \frac{1}{2} \int \mathrm{d}\boldsymbol{x}_1\, \mathrm{d}\boldsymbol{x}_2\, u(|\boldsymbol{x}_1 - \boldsymbol{x}_2|)\ \rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) \\
&= \frac{1}{2} \int \mathrm{d}\boldsymbol{x}_1 \int \mathrm{d}\boldsymbol{r}\, u(r)\, \rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_1 + \boldsymbol{r}) = \frac{V\rho^2}{2} \int \mathrm{d}\boldsymbol{r}\, u(r)\, g(r) \\
&= \frac{V\rho^2}{2} \int_0^\infty \mathrm{d}r\, 4\pi r^2\, u(r)\, g(r),
\end{aligned} \tag{7.23}$$

where we have set $r = |\boldsymbol{r}|$, and have exploited the invariance of $\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ with respect to translations, the definition of $g(r)$, and the isotropy of the system.

In the same hypotheses, the pressure can be expressed in terms of the $g(r)$, too:

$$\frac{p}{k_B T} = \rho - \frac{\rho^2}{2dk_B T} \int \mathrm{d}\boldsymbol{r}\, u'(r)\, g(r), \tag{7.24}$$

where d is the dimensionality of the system, $\boldsymbol{r} \in \mathbb{R}^d$, and $u'(r) = \mathrm{d}u/\mathrm{d}r$. In order to obtain this relation, let us suppose that the system is contained in a cubic box of side L, so that $V = L^d$. Then the pressure is given by

$$\frac{p}{k_B T} = \frac{\partial \ln Z_N}{\partial V}\Big)_{T,N} = \frac{L}{d\,V}\, \frac{\partial \ln Q_N}{\partial L}\Big)_{T,N}. \tag{7.25}$$

On the other hand, if $v(\boldsymbol{x}) = v_{\mathcal{B}}(\boldsymbol{x})$, the configuration integral is expressed by

$$\begin{aligned} Q_N &= \int_0^L \mathrm{d}r_{11} \int_0^L \mathrm{d}r_{12} \cdots \int_0^L \mathrm{d}r_{Nd}\, \exp\left[-\frac{U\left(\boldsymbol{r}^N\right)}{k_B T}\right] \\ &= V^N \int_0^1 \mathrm{d}\xi_{11} \int_0^1 \mathrm{d}\xi_{12} \cdots \int_0^1 \mathrm{d}\xi_{Nd}\, \exp\left[-\frac{U(L\boldsymbol{\xi})}{k_B T}\right], \end{aligned} \tag{7.26}$$

where we have set $\boldsymbol{r}_i = (r_{i1}, \ldots, r_{id}) = L(\xi_{i1}, \ldots, \xi_{id}) = L\boldsymbol{\xi}_i$, and $\boldsymbol{\xi} = (\boldsymbol{\xi}_i)$. Therefore,

$$L\frac{\partial Q_N}{\partial L} = dN - \frac{V^N}{k_B T\, Q_N} \int_0^1 \mathrm{d}\xi_{11} \cdots \int_0^1 \mathrm{d}\xi_{Nd}\, L\frac{\partial U(L\boldsymbol{\xi})}{\partial L}\, \exp\left[-\frac{U(L\boldsymbol{\xi})}{k_B T}\right]. \tag{7.27}$$

Now,

$$\begin{aligned} L\frac{\partial U(L\boldsymbol{\xi})}{\partial L} &= \frac{1}{2}\sum_{i\neq j} \frac{\partial u(L\left|\boldsymbol{\xi}_i - \boldsymbol{\xi}_i\right|)}{\partial L} = \frac{1}{2}\sum_{i\neq j} L\left|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\right|\ \left.\frac{\mathrm{d}u(r)}{\mathrm{d}r}\right|_{r=L\left|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\right|} \\ &= \frac{1}{2}\sum_{i\neq j} r_{ij}\, \left.\frac{\mathrm{d}u(r)}{\mathrm{d}r}\right|_{r=r_{ij}}, \end{aligned} \tag{7.28}$$

where $r_{ij} = \left|\boldsymbol{r}_i - \boldsymbol{r}_j\right|$. We thus have

$$\begin{aligned} L\frac{\partial Q_N}{\partial L} &= d\,N - \frac{1}{k_B T\, Q_N} \int \mathrm{d}\boldsymbol{r}^N\, \frac{1}{2}\sum_{i\neq j} r_{ij} u'(r_{ij})\, \exp\left[-\frac{U\left(\boldsymbol{r}^N\right)}{k_B T}\right] \\ &= d\,N - \frac{1}{2k_B T} \int \mathrm{d}\boldsymbol{x}_1 \int \mathrm{d}\boldsymbol{x}_2\ |\boldsymbol{x}_1 - \boldsymbol{x}_2|\ u'(|\boldsymbol{x}_1 - \boldsymbol{x}_2|)\, \rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ &= d\,N - \frac{V}{2k_B T}\rho^2 \int \mathrm{d}\boldsymbol{r}\, r\, u'(r)\, g(r) \\ &= 3N - \frac{V}{2k_B T}\rho^2 \int \mathrm{d}r\, 4\pi r^3\, u'(r)\, g(r), \end{aligned} \tag{7.29}$$

where we have exploited the homogeneity and isotropy of the system, and have set $r = |\boldsymbol{x}|$ and $d = 3$ in the last equality. By substituting this result in eq. (7.25), we obtain eq. (7.24). This expression is called the **pressure equation**, and we revisit it in sec. 8.2.1.

These expressions were obtained in the canonical ensemble, where the particle number N is fixed. However, $g(r)$ is independent of N, if N is large enough, since it is a

thermodynamic density, that is, the ratio of two extensive thermodynamic variables. We have in fact

$$\frac{N^2}{V} g(|\boldsymbol{x}|) = \int \mathrm{d}\boldsymbol{x}_0 \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(\boldsymbol{x}_0 - \boldsymbol{r}_i)\delta(\boldsymbol{x}_0 + \boldsymbol{x} - \boldsymbol{r}_j) \right\rangle. \tag{7.30}$$

The quantity on the right-hand side is an extensive thermodynamic variable. (It seems to be more than extensive, because it contains N^2 terms, but you can check and see that its value is actually proportional to the system size, since for each particle i there will be at most one particle j that will contribute to the sum.) By dividing it by the volume, we obtain a density, that is, a size-independent quantity. Thermodynamics guarantees us that all densities have the same value in a state of thermodynamic equilibrium, independently of the ensemble chosen to represent that state—therefore, the $g(r)$ has the same value in both the canonical and the grand canonical ensembles. We can use this relation to relate the $g(r)$ to the fluctuations in the number of particles in the grand canonical ensemble and, by means of the relation between fluctuations and susceptibility, to the isothermal compressibility $K_T = -V^{-1}\, \partial V/\partial p)_T$, as in the exercise that follows.

Exercise 7.2 Derive the **compressibility equation of state**:

$$\rho \int \mathrm{d}\boldsymbol{r}\ [g(r) - 1] + 1 = \rho\, k_\mathrm{B} T\, K_T.$$

We must, however, not fall into the error of supposing that all thermodynamic properties can be obtained from $g(r)$, not even for systems with only pair interactions and isotropic potentials. Indeed, knowing $g(r)$ is akin to knowing the internal energy E as a function of the temperature T and the volume V—it is useful information, but incomplete. More specifically, there is no way of evaluating the system entropy starting from the $g(r)$. This situation is unfortunate because entropy is crucial to understanding the behavior of liquids. The configurational entropy of a gas is complete in the sense that the particles can be essentially anywhere. That of a solid is minimal, since the particles perform only small oscillations around their equilibrium positions. A liquid, instead, has a sort of restricted entropy, and $g(r)$ provides only a partial description of this situation. By knowing only $g(r)$, one is not able to predict the stability of the liquid phase as compared to that of the other possible phases. In structural glasses, which can be considered metastable phases, the $g(r)$ is very similar to that of liquids, but the particles have restricted motion as in a solid. If one only knows the $g(r)$, one cannot distinguish a liquid from a glass.

7.2.2 Reversible Work Theorem

The pair distribution $g(r)$ can be expressed in terms of the work $W_\mathrm{rev}(r)$ needed to bring two particles reversibly from infinity to a distance r from each other. One has

$$g(r) = \exp\left[-\frac{W_\mathrm{rev}(r)}{k_\mathrm{B} T}\right]. \tag{7.31}$$

This relation is known as the **reversible work theorem**.

To prove it, we start from the definition of $\rho^{(2)}$:

$$\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{Q} \int \mathrm{d}\boldsymbol{r}^N \, \exp\left[-\frac{U\left(\boldsymbol{r}^N\right)}{k_\mathrm{B} T} \right] \delta(\boldsymbol{x}_1 - \boldsymbol{r}_1) \delta(\boldsymbol{x}_2 - \boldsymbol{r}_2). \tag{7.32}$$

We then have

$$\begin{aligned} \frac{\partial \ln \rho^{(2)}}{\partial \boldsymbol{x}_2} = & -\frac{1}{k_\mathrm{B} T} \int \mathrm{d}\boldsymbol{r}^N \, \frac{\partial U\left(\boldsymbol{r}^N\right)}{\partial \boldsymbol{x}_2} \exp\left[-\frac{U\left(\boldsymbol{r}^N\right)}{k_\mathrm{B} T} \right] \delta(\boldsymbol{x}_1 - \boldsymbol{r}_1) \delta(\boldsymbol{x}_2 - \boldsymbol{r}_2) \\ & \times \left[\int \mathrm{d}\boldsymbol{r}^N \exp\left[-\frac{U(\boldsymbol{r}^N)}{k_\mathrm{B} T} \right] \delta(\boldsymbol{x}_1 - \boldsymbol{r}_1) \delta(\boldsymbol{x}2 - \boldsymbol{r}2) \right]^{-1} \\ = & \frac{1}{k_\mathrm{B} T} \left\langle \boldsymbol{F}\left(\boldsymbol{x}_2; \boldsymbol{x}_1, \boldsymbol{r}^{N-2}\right) \right\rangle, \end{aligned} \tag{7.33}$$

where $\boldsymbol{F}\left(\boldsymbol{x}_2; \boldsymbol{x}_1, \boldsymbol{r}^{N-2}\right)$ is the force that acts on a particle in $\boldsymbol{x}_2$ when a particle is placed in $\boldsymbol{x}_1$ and the other particles are in the configuration $\boldsymbol{r}^{N-2}$. The force we must apply to the particle in order to keep the particle in $\boldsymbol{x}_2$ is the opposite of this force. By integrating this relation between infinity and $\boldsymbol{x}_2$, and taking into account the fact that $\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) \to \rho^2$ for $|\boldsymbol{x}_1 - \boldsymbol{x}_2| \to \infty$, we obtain

$$\ln \rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \ln \rho^2 - \frac{1}{k_\mathrm{B} T} \int_{x_2}^{\infty} \mathrm{d}\boldsymbol{x}_2 \cdot \left\langle \boldsymbol{F}\left(\boldsymbol{x}_2; \boldsymbol{x}_1, \boldsymbol{r}^{N-2}\right) \right\rangle. \tag{7.34}$$

Now, if $|\boldsymbol{r}| = r$, one has

$$W_\mathrm{rev}(r) = -\int_{\infty}^{r} \mathrm{d}\boldsymbol{x}_2 \cdot \left\langle \boldsymbol{F}\left(\boldsymbol{x}_2; 0, \boldsymbol{r}^{N-2}\right) \right\rangle, \tag{7.35}$$

and $\rho^{(2)}(0, \boldsymbol{r}) = \rho^2 g(r)$. Therefore,

$$\ln g(r) = -\frac{1}{k_\mathrm{B} T} W_\mathrm{rev}(r). \tag{7.36}$$

7.2.3 Measurement of g(r)

The interest in $g(r)$ is enhanced by the fact that its Fourier transform is directly measurable by means of neutron scattering experiments.

Neutrons interact only with the particle nuclei, and the nuclei are so small that the interaction potential is well represented by a delta function. In a scattering experiment, one prepares a monochromatic beam of neutrons, characterized by the momentum $\boldsymbol{p}$, related to the wave vector $\boldsymbol{q}$ by the de Broglie relation

$$\boldsymbol{p} = \hbar \boldsymbol{q}. \tag{7.37}$$

The neutrons are produced in a nuclear reactor and are selected by subjecting them to Bragg elastic scattering, which selects the neutrons with the correct value of $\boldsymbol{q}$. The beam of monochromatic neutrons is then directed onto the sample, and one measures the flow of neutrons that have been scattered by a certain angle θ. The geometry of the experiment is shown in fig. 7.1. After the scattering, the neutrons are characterized by a wave vector

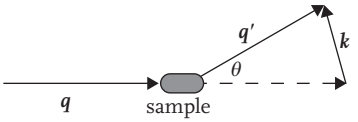


FIGURE 7.1. Scheme of a neutron scattering experiment. The incoming neutron flux with wave vector $\boldsymbol{q}$ is elastically scattered by the sample, acquiring the wave vector $\boldsymbol{q}'$, with $|\boldsymbol{q}'| = |\boldsymbol{q}| = q$. The angle between $\boldsymbol{q}'$ and $\boldsymbol{q}$ is θ, and we have $\boldsymbol{q}' = \boldsymbol{q} + \boldsymbol{k}$, with $|\boldsymbol{k}| = 2q\sin(\theta/2)$.

$\boldsymbol{q}'$, whose magnitude is equal to that of $\boldsymbol{q}$, because the scattering is elastic. The result of the experiment therefore, is the ratio between the outgoing flux, with a certain wave vector $\boldsymbol{q}'$, and the incoming flux, as a function of the difference wave vector $\boldsymbol{k}$. For an isotropic system, the intensity is independent of the direction of $\boldsymbol{k}$ and, as one can easily derive from the geometry shown in the illustration, one has

$$|\boldsymbol{k}| = 2q\sin\frac{\theta}{2}, \tag{7.38}$$

where $q = |\boldsymbol{q}| = |\boldsymbol{q}'|$.

We now discuss how one can obtain a direct measurement of the structure function $S(\boldsymbol{k})$ from this experiment. A wave function $\psi_{\boldsymbol{q}}(\boldsymbol{x}) = \psi_0\, \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}}$ that interacts with a nucleus placed in $\boldsymbol{r}_i$ produces, in each direction determined by the wave vector $\boldsymbol{q}'$, a diffused wave:

$$\psi'_{\boldsymbol{k}}(\boldsymbol{x}) = \psi_0\, \phi(\boldsymbol{k})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}} \frac{\mathrm{e}^{-\mathrm{i}\boldsymbol{q}'\cdot(\boldsymbol{x}-\boldsymbol{r}_i)}}{|\boldsymbol{x}-\boldsymbol{r}_i|}, \tag{7.39}$$

where the coefficient $\phi(\boldsymbol{k})$ is the scattering amplitude expressed by the matrix element of the interaction potential $u_{\mathrm{nn}}(\boldsymbol{r})$ between neutron and nucleus,

$$\phi(\boldsymbol{k}) = \langle \boldsymbol{q}' |\, u_{\mathrm{nn}}\, | \boldsymbol{q} \rangle \propto \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{q}'\cdot\boldsymbol{r}+\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}} u_{\mathrm{nn}}(\boldsymbol{r}) \propto u_{\mathrm{nn}}(\boldsymbol{k}). \tag{7.40}$$

In the case we are considering, in which the potential is well approximated by a delta, $\phi(\boldsymbol{k})$ is approximated by a constant ϕ_0. In general, $\phi(\boldsymbol{k})$ will have a nontrivial dependence on $\boldsymbol{k}$ for wave vectors on the order of the inverse of the characteristic lengths of the scattering center. Since we are interested in much greater distances (and consequently in much smaller values of $|\boldsymbol{k}|$), we can disregard this dependence.

The wave function measured by the detector at a point $\boldsymbol{R}$ is the superposition of the wave functions scattered by the individual scattering centers:

$$\Psi(\boldsymbol{R}) = \psi_0\phi_0 \sum_i \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_i} \frac{\mathrm{e}^{-\mathrm{i}\boldsymbol{q}'\cdot(\boldsymbol{R}-\boldsymbol{r}_i)}}{|\boldsymbol{R}-\boldsymbol{r}_i|}. \tag{7.41}$$

If we assume that the sample is small with respect to $|\boldsymbol{R}|$ and is placed at the origin, one has $|\boldsymbol{R}-\boldsymbol{r}_i| \simeq |\boldsymbol{R}|$, and we can disregard the dependence on $\boldsymbol{r}_i$ in the denominator. The flux of neutrons in $\boldsymbol{R}$ is proportional to the square modulus of $\Psi(\boldsymbol{R})$:

$$\left|\Psi(\boldsymbol{R})\right|^2 = |\psi_0|^2 \frac{|\phi_0|^2}{R^2} \sum_{ij} \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_i - \mathrm{i}\boldsymbol{q}'(\boldsymbol{R}-\boldsymbol{r}_i)} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}_j + \mathrm{i}\boldsymbol{q}'\cdot(\boldsymbol{R}-\boldsymbol{r}_j)}. \tag{7.42}$$

The dependence on $\boldsymbol{R}$ in the exponent cancels out, and one remains with

$$\left|\Psi(\boldsymbol{R})\right|^2 = |\psi_0|^2 \frac{|\phi_0|^2}{R^2} \sum_{ij} \mathrm{e}^{\mathrm{i}(\boldsymbol{q}-\boldsymbol{q}')\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)}. \tag{7.43}$$

From this quantity, we isolate the terms in which $i=j$. We obtain

$$\left|\Psi(\boldsymbol{R})\right|^2 = |\psi_0|^2 \frac{|\phi_0|^2}{R^2} \left[N + \sum_{i\neq j} \mathrm{e}^{\mathrm{i}(\boldsymbol{q}-\boldsymbol{q}')\cdot(\boldsymbol{r}_i-\boldsymbol{r}_j)} \right]. \tag{7.44}$$

Taking into account the fact that $\boldsymbol{k}=\boldsymbol{q}'-\boldsymbol{q}$, and evaluating the average of this expression with respect to thermal fluctuations, we obtain

$$\begin{aligned}
\left\langle \left|\Psi(\boldsymbol{R})\right|^2 \right\rangle &= |\psi_0|^2 \frac{|\phi_0|^2}{R^2} \left[N + \int \mathrm{d}\boldsymbol{x} \int \mathrm{d}\boldsymbol{x}' \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} \left\langle \sum_{i\neq j} \delta(\boldsymbol{x}-\boldsymbol{r}_i)\delta(\boldsymbol{x}'-\boldsymbol{r}_j) \right\rangle \right] \\
&= |\psi_0|^2 \frac{|\phi_0|^2}{R^2} \left[N + \int \mathrm{d}\boldsymbol{x} \int \mathrm{d}\boldsymbol{x}' \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} \rho^{(2)}(\boldsymbol{x},\boldsymbol{x}') \right] \\
&= |\psi_0|^2 \frac{|\phi_0|^2}{R^2} N \left[1 + \frac{N}{V} \int \mathrm{d}\boldsymbol{r} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} g(r) \right].
\end{aligned} \tag{7.45}$$

By evaluating the ratio between the outgoing and the incoming particle fluxes; the factor $|\psi_0|^2$ cancels out. On the other hand, by evaluating the outgoing flux in a solid angle element $\mathrm{d}\Omega = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi$, the area over which we integrate the flux becomes proportional to $|\boldsymbol{R}|^2$, which cancels out with the denominator. Since for $\boldsymbol{k}=0$, moreover, there is no way of distinguishing between forward-scattered particles and nonscattered particles, there is no reason to also use this formula for $\boldsymbol{k}=0$. It is usual to subtract a delta function at $\boldsymbol{k}=0$ from this expression, thus obtaining the following expression of the ratio between entrance and exit fluxes in the direction defined by $\boldsymbol{q}'$:

$$\frac{I(\theta)}{I_0} = |\phi_0|^2 \, N \left[1 + \rho \int \mathrm{d}\boldsymbol{r} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \, [g(r)-1] \right]. \tag{7.46}$$

We can now define the **structure factor** $S(\boldsymbol{k})$:

$$S(\boldsymbol{k}) = 1 + \rho \int \mathrm{d}\boldsymbol{r} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \, [g(r)-1]. \tag{7.47}$$

Then, for $\boldsymbol{k}\neq 0$, one has

$$I(\theta) = I_0 \, |\phi_0|^2 \, S(\boldsymbol{k}). \tag{7.48}$$

By comparing with eq. (7.18) we see that $S(\boldsymbol{k})$ is the Fourier transform of $G(0,\boldsymbol{x})/\rho$:

$$\int \mathrm{d}\boldsymbol{r} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \, G(0,\boldsymbol{r}) = \rho \int \mathrm{d}\boldsymbol{r} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \left\{ 1 + \rho[g(r)-1] \right\} = \rho \, S(\boldsymbol{k}). \tag{7.49}$$

Measuring the outgoing flux therefore provides a measure of $S(\boldsymbol{k})$, which in turn allows us to determine the pair distribution $g(r)$ by means of an inverse Fourier transformation.

7.2.4 Direct Correlation Function

The $g(r)$ is connected to the isothermal compressibility by the compressibility equation of state, which we obtained in exercise 7.1. We know that, since the correlation length diverges at the critical point, the compressibility diverges by approaching it, too. Ornstein and Zernicke had the idea of separating the effects that lead to the divergence of the correlation length in the $g(r)$, by introducing the direct correlation function $c(r)$, which is always short-range, and which one expects to be regular around the critical point. Actually, we know from the theory of critical phenomena that this is not exactly true, and can be so only in the framework of a mean-field theory.

It is useful to introduce the function $h(r) = g(r) - 1$, which goes to zero for $r \to \infty$. The **direct correlation function** $c(r)$ is then defined by

$$h(r) = c(r) + \rho \int \mathrm{d}\boldsymbol{r}'\, c(r')h(|\boldsymbol{r} - \boldsymbol{r}'|). \tag{7.50}$$

The correlations expressed by $h(r)$ are thus divided into two contributions: the *direct correlations* between particles placed at a distance r, represented by $c(r)$, and the indirect correlations, in which the particle at the origin directly influences a particle placed at $\boldsymbol{r}'$, which in its turn (directly or indirectly) influences the particle placed at $\boldsymbol{r}$.

This integral equation can be solved by a Fourier transformation:

$$h(\boldsymbol{k}) = c(\boldsymbol{k}) + \rho\, c(\boldsymbol{k})\, h(\boldsymbol{k}). \tag{7.51}$$

We thus obtain

$$h(\boldsymbol{k}) = \frac{c(\boldsymbol{k})}{1 - \rho\, c(\boldsymbol{k})}. \tag{7.52}$$

The divergence of isothermal compressibility is exhibited by the divergence of $h(\boldsymbol{k})$ for $|\boldsymbol{k}| \to 0$. From this equation, we see that this occurs if $c(\boldsymbol{k}{=}0) \to 1/\rho$ for $T \to T_c$.

These equations obviously do not provide us with any new information, if we are not able to evaluate the $h(r)$ or the $c(r)$ independently. We will see in the following sections that various approximation schemes can be formulated in terms of an additional relation between the $h(r)$ (and therefore the $g(r)$) and the $c(r)$. One thus has a closed set of equations, which leads to reasonable results when the density is not too large.

7.2.5 The BBGKY Hierarchy

The various reduced densities are connected by a chain of equations, so that each $\rho^{(n)}$ can be expressed as a function of $\rho^{(n+1)}$ and of the interaction potential. We are therefore dealing with an equation hierarchy, which has been given the initials of the many authors who discovered it (actually, in the time-dependent version): Born, Bogolyubov, Green, Kirkwood, Yvon. It is therefore referred to as the **BBGKY hierarchy**.

Let us consider the derivative of $\rho^{(1)}(\boldsymbol{x}_1)$ with respect to $\boldsymbol{x}_1$:

$$\begin{aligned}\frac{\partial \rho^{(1)}}{\partial \boldsymbol{x}_1} &= \frac{N}{Q_N}\frac{\partial}{\partial \boldsymbol{x}_1}\int \mathrm{d}\boldsymbol{r}^N \exp\left[-\frac{U\left(\boldsymbol{r}^N\right)}{k_\mathrm{B}T}\right]\delta(\boldsymbol{x}_1-\boldsymbol{r}_1)\\ &= -\frac{N}{k_\mathrm{B}TQ_N}\int \mathrm{d}\boldsymbol{r}^N \left(\frac{\partial}{\partial \boldsymbol{x}_1}\sum_{j(\neq 1)} u(|\boldsymbol{x}_1-\boldsymbol{r}_j|)\right)\\ &\qquad \times \exp\left[-\frac{U\left(\boldsymbol{r}^N\right)}{k_\mathrm{B}T}\right]\delta(\boldsymbol{x}_1-\boldsymbol{r}_1)\\ &= -\frac{N(N-1)}{k_\mathrm{B}TQ_N}\int \mathrm{d}\boldsymbol{x}_2 \left(\frac{\partial}{\partial \boldsymbol{x}_1}u(|\boldsymbol{x}_1-\boldsymbol{x}_2|)\right)\\ &\qquad \times \int \mathrm{d}\boldsymbol{r}^N \exp\left[-\frac{U\left(\boldsymbol{r}^N\right)}{k_\mathrm{B}T}\right]\delta(\boldsymbol{x}_1-\boldsymbol{r}_1)\delta(\boldsymbol{x}_2-\boldsymbol{r}_2)\\ &= -\frac{1}{k_\mathrm{B}T}\int \mathrm{d}\boldsymbol{x}_2 \left(\frac{\partial}{\partial \boldsymbol{x}_1}u(|\boldsymbol{x}_1-\boldsymbol{x}_2|)\right)\rho^{(2)}(\boldsymbol{x}_1,\boldsymbol{x}_2).\end{aligned} \tag{7.53}$$

In the same vein, one obtains a relation expressing the derivative of $\rho^{(2)}(\boldsymbol{x}_1,\boldsymbol{x}_2)$ in terms of $\rho^{(3)}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)$:

$$\begin{aligned}\frac{\partial \rho^{(2)}}{\partial \boldsymbol{x}_1} &= -\frac{1}{k_\mathrm{B}T}\left(\frac{\partial}{\partial \boldsymbol{x}_1}u(|\boldsymbol{x}_1-\boldsymbol{x}_2|)\right)\rho^{(2)}(\boldsymbol{x}_1-\boldsymbol{x}_2)\\ &\quad -\frac{1}{k_\mathrm{B}T}\int \mathrm{d}\boldsymbol{x}_3 \left(\frac{\partial}{\partial \boldsymbol{x}_1}u(|\boldsymbol{x}_1-\boldsymbol{x}_2|)\right)\rho^{(3)}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3).\end{aligned} \tag{7.54}$$

Let us introduce the generalizations of the pair distribution function $g(r)$:

$$\rho^{(2)}(\boldsymbol{x}_1,\boldsymbol{x}_2) = \rho^2\, g(\boldsymbol{x}_1,\boldsymbol{x}_2), \tag{7.55}$$

$$\rho^{(3)}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3) = \rho^3\, g^{(3)}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3). \tag{7.56}$$

We thus obtain

$$\begin{aligned}\frac{\partial g}{\partial \boldsymbol{x}_1} &= -\frac{1}{k_\mathrm{B}T}\left(\frac{\partial}{\partial \boldsymbol{x}_1}u(|\boldsymbol{x}_1-\boldsymbol{x}_2|)\right)g(\boldsymbol{x}_1-\boldsymbol{x}_2)\\ &\quad -\frac{\rho}{k_\mathrm{B}T}\int \mathrm{d}\boldsymbol{x}_3 \left(\frac{\partial}{\partial \boldsymbol{x}_1}u(|\boldsymbol{x}_1-\boldsymbol{x}_3|)\right)g^{(3)}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3).\end{aligned} \tag{7.57}$$

Clearly, the same method allows one to express the derivative of the reduced density of an arbitrary order n in terms of expressions involving reduced densities of order up to $(n+1)$. Obviously, there is no way to solve directly these equations, but it is possible to use this hierarchy as a starting point for some approximations, such as Kirkwood's **superposition approximation**. It consists of the representation of $g^{(3)}$ in terms of a product of pair distributions:

$$g^{(3)}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3) = g(\boldsymbol{x}_1,\boldsymbol{x}_2)\, g(\boldsymbol{x}_2,\boldsymbol{x}_3)\, g(\boldsymbol{x}_3,\boldsymbol{x}_1). \tag{7.58}$$

This approximation allows one to convert eq. (7.57) into a closed nonlinear equation for the $g(r)$, which is known as the Born-Green-Yvon (BGY) equation, and turns out to be fairly useful at low densities.

Indeed, by introducing the shorthand $r_{ij} = |\boldsymbol{x}_i - \boldsymbol{x}_j|$, and substituting this approximation in (7.57), we obtain

$$\begin{aligned}\frac{\partial g}{\partial \boldsymbol{x}_1} = &-\frac{1}{k_B T}\left(\frac{\partial}{\partial \boldsymbol{x}_1} u(r_{12})\right) g(r_{12}) \\ &-\frac{\rho}{k_B T}\int \mathrm{d}\boldsymbol{x}_3 \left(\frac{\partial}{\partial \boldsymbol{x}_1} u(r_{13})\right) g(r_{12})\, g(r_{23})\, g(r_{31}),\end{aligned} \tag{7.59}$$

which can be rearranged to yield

$$\frac{\partial}{\partial \boldsymbol{x}_1}\left[\ln g(r_{12}) + \frac{u(r_{12})}{k_B T}\right] = -\frac{\rho}{k_B T}\int \mathrm{d}\boldsymbol{x}_3 \left(\frac{\partial}{\partial \boldsymbol{x}_1} u(r_{13})\right) g(r_{23})\, g(r_{31}). \tag{7.60}$$

This is the BGY equation, and it can also be expressed in another form, which was derived by Kirkwood. Let us introduce the auxiliary function

$$E(r) = -\int_r^\infty \mathrm{d}r' \, \frac{\mathrm{d}u(r')}{\mathrm{d}r'}\, g(r'). \tag{7.61}$$

We then have

$$\frac{\partial E(r_{13})}{\partial \boldsymbol{x}_1} = \frac{\partial u(r_{13})}{\partial \boldsymbol{x}_1} g(r_{13}). \tag{7.62}$$

By substituting this expression in (7.60), we obtain

$$\frac{\partial}{\partial \boldsymbol{x}_1}\left[\ln g(r_{12}) + \frac{u(r_{12})}{k_B T}\right] = -\frac{\rho}{k_B T}\int \mathrm{d}\boldsymbol{x}_3 \, \frac{\partial E(r_{13})}{\partial \boldsymbol{x}_1}\, g(r_{23}). \tag{7.63}$$

By integrating with respect to $\boldsymbol{x}_1$, we obtain

$$\ln g(r_{12}) + \frac{u(r_{12})}{k_B T} = -\frac{\rho}{k_B T}\int \mathrm{d}\boldsymbol{x}_3 \, E(r_{13}) g(r_{23}) + C, \tag{7.64}$$

where C is an integration constant. We can evaluate C by going to the limit $r_{12} \to \infty$, keeping r_{13} finite. The left-hand side vanishes, and $g(r_{23}) \to 1$. We thus obtain

$$C = \frac{\rho}{k_B T}\int \mathrm{d}\boldsymbol{r}\, E(r). \tag{7.65}$$

By substituting this result in (7.64), we obtain

$$\ln g(r_{12}) + \frac{u(r_{12})}{k_B T} = -\frac{\rho}{k_B T}\int \mathrm{d}\boldsymbol{x}_3 \, E(r_{13})\,[g(r_{23}) - 1], \tag{7.66}$$

a result known as the **Kirkwood equation**. A much more complete discussion of the BBGKY hierarchy is contained in [Bark76].

7.3 The Density Functional Approach

In a system described by the grand canonical ensemble, the thermodynamic properties are obtained from the thermodynamic potential Ω. This is a function of μ, T, and the volume V, which satisfies the relations

$$\left.\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = -N; \qquad \left.\frac{\partial\Omega}{\partial T}\right)_{\mu,V} = -S; \qquad \left.\frac{\partial\Omega}{\partial V}\right)_{\mu,T} = -p. \tag{7.67}$$

It is related to the grand partition function Z_{GC} by

$$\Omega = -k_{\mathrm{B}}T \ln Z_{\mathrm{GC}}. \tag{7.68}$$

When the system is subject to a nonuniform potential $v(\boldsymbol{x})$, the partition function Z_{GC} becomes a functional of v. Considering the expression of the grand partition function

$$Z_{\mathrm{GC}} = \sum_{N=0}^{\infty} \frac{1}{N!\lambda_{\mathrm{B}}^{3N}} \int \mathrm{d}\boldsymbol{r}^N \exp\left[-\frac{1}{k_{\mathrm{B}}T}\left(U\left(\boldsymbol{r}^N\right) + \sum_{i=1}^{N}\left(v(\boldsymbol{r}_i) - \mu\right)\right)\right], \tag{7.69}$$

where $U(\boldsymbol{r}^N) = U(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$, we see that Z_{GC}, and therefore Ω, can be considered functionals of the "local" chemical potential

$$\psi(\boldsymbol{x}) = \mu - v(\boldsymbol{x}). \tag{7.70}$$

Now

$$\sum_{i=1}^{N} \psi(\boldsymbol{r}_i) = \int \mathrm{d}\boldsymbol{r} \left(\sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i)\right) \psi(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}\, \hat{\rho}_N(\boldsymbol{r})\psi(\boldsymbol{r}), \tag{7.71}$$

where $\hat{\rho}_N(\boldsymbol{x})$ is the fluctuating local density at point $\boldsymbol{x}$. Therefore Z_{GC} can be expressed by

$$Z_{\mathrm{GC}} = \sum_{N=0}^{\infty} \frac{1}{N!\lambda_{\mathrm{B}}^{3N}} \int \mathrm{d}\boldsymbol{r}^N \exp\left[-\frac{1}{k_{\mathrm{B}}T}\left(U\left(\boldsymbol{r}^N\right) - \int \mathrm{d}\boldsymbol{r}\, \hat{\rho}_N(\boldsymbol{r})\psi(\boldsymbol{r})\right)\right]. \tag{7.72}$$

We can now straightforwardly evaluate the functional derivative of $\ln Z_{\mathrm{GC}}$ with respect to $\psi(\boldsymbol{x})$:

$$\begin{aligned}\frac{\delta \ln Z_{\mathrm{GC}}}{\delta\psi(\boldsymbol{x})} &= \frac{1}{Z_{\mathrm{GC}}}\frac{1}{k_{\mathrm{B}}T}\sum_{N=0}^{\infty}\frac{1}{N!\lambda_{\mathrm{B}}^{3N}}\int \mathrm{d}\boldsymbol{r}^N \\ &\qquad \times \exp\left[-\frac{1}{k_{\mathrm{B}}T}\left(U\left(\boldsymbol{r}^N\right) - \int \mathrm{d}\boldsymbol{r}\, \hat{\rho}_N(\boldsymbol{r})\psi(\boldsymbol{r})\right)\right]\hat{\rho}_N(\boldsymbol{r}) \\ &= \frac{1}{k_{\mathrm{B}}T}\rho(\boldsymbol{x}),\end{aligned} \tag{7.73}$$

where we have exploited the relation

$$\rho(\boldsymbol{x}) = \left\langle \hat{\rho}_N(\boldsymbol{x}) \right\rangle, \tag{7.74}$$

in which the average is taken in the grand canonical ensemble. We have, therefore, taking into account eq. (7.68),

$$\frac{\delta\Omega}{\delta\psi(\boldsymbol{x})} = -\rho(\boldsymbol{x}). \tag{7.75}$$

Importantly, it may be shown that while the function $\rho(\boldsymbol{x})$ is uniquely identified by the function $\psi(\boldsymbol{x})$ via eq. (7.75), it is also true that the function $\psi(\boldsymbol{x})$ is uniquely identified by $\rho(\boldsymbol{x})$. We may therefore generalize the Legendre transformation (2.105) connecting Ω with the Helmholtz free energy F in thermodynamics to the case of nonuniform systems, via the relation

$$\Omega[\psi] = F[\rho] - \int \mathrm{d}\boldsymbol{x}\, \rho(\boldsymbol{x})\psi(\boldsymbol{x}), \tag{7.76}$$

where $F[\rho]$ is a functional of the local density $\rho = (\rho(\boldsymbol{x}))$ and the function $\psi = (\psi(\boldsymbol{x}))$ is determined by solving eq. (7.68) with the given ρ. Thus the function ρ appears as the conjugate to the function ψ. Now, all explicit dependence on the single-particle potential $v(\boldsymbol{x})$ has disappeared in $F[\rho]$, which can therefore be considered as an intrinsic property of the fluid.

The free energy functional satisfies the variational principle of the Legendre transform. We have

$$F[\rho] - \int \mathrm{d}\boldsymbol{x}\, \rho(\boldsymbol{x})\psi(\boldsymbol{x}) \geq \Omega[\psi], \tag{7.77}$$

for all local density distributions ρ, and the equality holds only for the distribution that corresponds to the equilibrium for the given ψ. This implies that the equilibrium density must satisfy the equation

$$\frac{\delta F[\rho]}{\delta\rho(\boldsymbol{x})} + v(\boldsymbol{x}) = \mu, \qquad \forall \boldsymbol{x} \in \mathcal{B}. \tag{7.78}$$

Exercise 7.3 Evaluate $F[\rho]$ for the case of a fluid of noninteracting particles and show that this equation yields the Boltzmann distribution in an external potential $v(\boldsymbol{x})$.

We owe the density functional approach to Hohenberg and Kohn [Hohe64], who introduced it to treat the zero-temperature inhomogeneous electron gas. It was then extended to positive temperatures by Mermin [Merm65].

7.3.1 Approximations to $F[\rho]$

Except for the case of noninteracting particles, it is not possible to evaluate $F[\rho]$ exactly, although some approximate expressions are available.

Let us consider the situation where the inhomogeneity $\delta\rho(\boldsymbol{x})$ is small with respect to the average uniform density $\rho_0 = N/V$. In the absence of an external perturbation, the free energy must be a minimum at the equilibrium value of the density ρ_0. We can then

expand $F[\rho]$ to quadratic order in $\delta\rho(\boldsymbol{x})$, obtaining

$$F[\rho_0 + \delta\rho(\boldsymbol{x})] \simeq F[\rho_0] + \frac{1}{2}\int \mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{x}'\, K(\boldsymbol{x},\boldsymbol{x}')\,\delta\rho(\boldsymbol{x})\,\delta\rho(\boldsymbol{x}'). \tag{7.79}$$

Here $F[\rho_0]$ is the free energy of the uniform system, and can therefore be written in terms of the equilibrium free energy density f_0 and the volume of the system as Vf_0. The function $K(\boldsymbol{x},\boldsymbol{x}')$ is the second functional derivative of $F[\rho]$, evaluated for $\rho(\boldsymbol{x}) = \rho_0$, $\forall \boldsymbol{x}$. Expressing eq. (7.79) in terms of the Fourier transform $\delta\rho(\boldsymbol{k})$ of $\delta\rho(\boldsymbol{x})$, we obtain

$$F[\rho_0 + \delta\rho] \simeq Vf_0 + \frac{1}{2}\sum_{\boldsymbol{k}} K(\boldsymbol{k})\,\delta\rho(\boldsymbol{k})\,\delta\rho(-\boldsymbol{k}). \tag{7.80}$$

By exploiting eq. (7.78), we can use this equation to show that, for a small nonuniform external potential $v(\boldsymbol{x})$, one has

$$K(\boldsymbol{k})\,\delta\rho(\boldsymbol{k}) = -v(\boldsymbol{k}). \tag{7.81}$$

Therefore $1/K(\boldsymbol{k})$ is the Fourier transform of the generalized susceptibility $\chi(\boldsymbol{x},\boldsymbol{x}')$, defined by

$$\delta\rho(\boldsymbol{x}) = \int \mathrm{d}\boldsymbol{x}'\,\chi(\boldsymbol{x},\boldsymbol{x}')\,\delta\psi(\boldsymbol{x}'). \tag{7.82}$$

By use of eqs. (7.75) and (7.73), we obtain, evaluating the second functional derivative,

$$\begin{aligned}\chi(\boldsymbol{x},\boldsymbol{x}') &= \frac{\delta\rho(\boldsymbol{x})}{\delta\psi(\boldsymbol{x}')} = -\frac{\delta^2\Omega}{\delta\psi(\boldsymbol{x})\,\delta\psi(\boldsymbol{x}')} \\ &= \frac{1}{k_{\mathrm{B}}T}\left[\rho_0\delta(\boldsymbol{x}-\boldsymbol{x}') + \rho^{(2)}(\boldsymbol{x},\boldsymbol{x}') - \rho_0^2\right] \\ &= \frac{\rho_0}{k_{\mathrm{B}}T}\left\{\delta(\boldsymbol{x}-\boldsymbol{x}') + \rho_0\left[g(|\boldsymbol{x}-\boldsymbol{x}'|) - 1\right]\right\}.\end{aligned} \tag{7.83}$$

Taking the Fourier transform of this relation, we obtain

$$\chi(\boldsymbol{k}) = \frac{\rho_0}{k_{\mathrm{B}}T}S(\boldsymbol{k}), \tag{7.84}$$

where $S(\boldsymbol{k})$ is the structure factor defined in eq. (7.47). Summing up, for small deviations $\delta\rho(\boldsymbol{k})$ from the uniform density ρ_0, one has

$$\delta F[\rho_0 + \delta\rho] \simeq \frac{k_{\mathrm{B}}T}{V}\sum_{\boldsymbol{k}}\frac{\delta\rho(\boldsymbol{k})\,\delta\rho(-\boldsymbol{k})}{S(\boldsymbol{k})}, \tag{7.85}$$

up to terms of higher order in $\delta\rho$.

If the inhomogeneity is slowly varying, that is, if $\nabla\rho/\rho \ll \xi^{-1}$, where ξ is the typical range of correlations in the uniform fluid (for example, the range of oscillations of $g(r)$), one can attempt to write $F[\rho]$ in the form

$$F[\rho] = \int \mathrm{d}\boldsymbol{x}\,\mathcal{F}\left(\rho(\boldsymbol{x}), \nabla\rho(\boldsymbol{x}), \nabla^2\rho(\boldsymbol{x}), \ldots\right), \tag{7.86}$$

where $\mathcal{F}$ is a function of its arguments, which are all evaluated at the running point $\boldsymbol{x}$. To lowest order, one has

$$F[\rho] = \int \mathrm{d}\boldsymbol{x} \left[f_0(\rho(\boldsymbol{x})) + \frac{1}{2}\kappa\,(\rho(\boldsymbol{x}))\,(\nabla\rho(\boldsymbol{x}))^2 + \cdots \right]. \tag{7.87}$$

If we neglect all terms in brackets beyond the first, we have the local-density approximation (LDA), which was famously used by Einstein in his 1910 work on critical opalescence [Eins10]. It consists in evaluating the free energy of the inhomogeneous system by summing up the contributions of different, nearly homogeneous, regions. Symmetry dictates that the first nontrivial term in $\nabla\rho$ be quadratic. Equation (7.87) is called the **square gradient** approximation.

If the inhomogeneity is both small and slowly varying, we have, in Fourier space,

$$F[\rho] \simeq V \left[f_0(\rho_0) + \frac{1}{2}\sum_{\boldsymbol{k}} \left[K_T^{-1} + \kappa(\rho_0)k^2 \right] \delta\rho(\boldsymbol{k})\,\delta\rho(-\boldsymbol{k}) \right], \tag{7.88}$$

where K_T is the isothermal compressibility and $\kappa(\rho_0)$ is a positive quantity. In this approximation, the correlation lengthL ξ of the density fluctuations is therefore given by

$$\xi = \sqrt{\kappa(\rho_0)K_T}. \tag{7.89}$$

7.3.2 Fluid-Fluid Interfaces

Let us consider a binary mixture on its phase-coexistence line. We assume that the interface between the A-rich and the B-rich phases is horizontal. At equilibrium, the total density $\rho_{\mathrm{tot}}(\boldsymbol{x}) = \rho_{\mathrm{A}}(\boldsymbol{x}) + \rho_{\mathrm{B}}(\boldsymbol{x})$ is fixed by the local composition, the pressure, and the temperature. Hence we can choose the density of one component, for example, A, as the variable of interest, and denote it simply by $\rho(x)$. We exploit translation invariance in the xy plane and we use the square gradient approximation to the free energy of the interface in the form

$$F[\rho] \simeq L^2 \int \mathrm{d}z \left[f(\rho(z)) + \frac{1}{2}\kappa\,(\rho(z)) \left(\frac{\mathrm{d}\rho(z)}{\mathrm{d}z} \right)^2 \right] = L^2 \int \mathrm{d}z\, \mathcal{L}\,(\rho, \dot{\rho}), \tag{7.90}$$

where L^2 is the interface area and we have set $\dot{\rho} = \mathrm{d}\rho/\mathrm{d}z$. This expression looks like the action integral in Lagrangian mechanics, with $-f(\rho)$ playing the role of the potential energy and $\kappa(\rho)$ that of a (ρ-dependent) mass. We thus obtain the "equation of motion" satisfied by the minimum free-energy profile:

$$\frac{\mathrm{d}}{\mathrm{d}z}\frac{\partial\mathcal{L}}{\partial\dot{\rho}} = \frac{\partial\mathcal{L}}{\partial\rho}. \tag{7.91}$$

This leads to the differential equation

$$\kappa(\rho)\frac{\mathrm{d}^2\rho}{\mathrm{d}z^2} + \frac{1}{2}\frac{\mathrm{d}\kappa}{\mathrm{d}\rho}\left(\frac{\mathrm{d}\rho}{\mathrm{d}z}\right)^2 = \frac{\mathrm{d}f}{\mathrm{d}\rho}. \tag{7.92}$$

By multiplying both sides by $\mathrm{d}\rho = \mathrm{d}z$, and integrating, we obtain (in analogy with energy conservation)

$$\frac{1}{2}\kappa(\rho)\left(\frac{\mathrm{d}\rho}{\mathrm{d}z}\right)^2 - f(\rho) = \text{const.} \tag{7.93}$$

The boundary conditions require, for example, $\lim_{z\to-\infty}\rho(z) = \rho^{(1)}$ and $\lim_{z\to+\infty}\rho(z) = \rho^{(2)}$, where $\rho^{(1)}$ and $\rho^{(2)}$ are the equilibrium densities of species A in the two coexisting phases. This requires the constant to be equal to the common equilibrium value of $f(\rho)$ in the two phases, which we can set to zero. We can solve this equation for $\mathrm{d}\rho/\mathrm{d}z$, insert the resulting expression in eq. (7.90), change the integration variable from z to ρ (choosing the correct sign for $\mathrm{d}\rho/\mathrm{d}z$), and obtain the excess free energy due to the presence of the interface:

$$\Delta F = L^2 \int_{\rho^{(1)}}^{\rho^{(2)}} \mathrm{d}\rho\, \sqrt{2\kappa(\rho)f(\rho)}, \tag{7.94}$$

where the integral goes from the smaller to the larger between $\rho^{(1)}$ and $\rho^{(2)}$.

To proceed, we need an expression for $f(\rho)$. Since $f(\rho)$ must exhibit minima at both $\rho^{(1)}$ and $\rho^{(2)}$, we introduce a simple quartic approximation of the form

$$f(\rho) = \frac{1}{2}C\left(\rho-\rho^{(1)}\right)^2\left(\rho-\rho^{(2)}\right)^2. \tag{7.95}$$

Supposing $\kappa(\rho) = \kappa = \text{const.}$ and substituting in (7.94), we obtain the surface tension γ:

$$\gamma = \frac{\Delta F}{L^2} = \sqrt{\kappa C}\left|\rho^{(2)}-\rho^{(1)}\right|^3. \tag{7.96}$$

In the Landau theory, $\left|\rho^{(2)}-\rho^{(1)}\right| \sim |T_\mathrm{c}-T|^{1/2}$. We thus obtain $\gamma \sim |T_\mathrm{c}-T|^{3/2}$ as the mean-field prediction. From scaling, we obtain instead $\gamma \sim |T_\mathrm{c}-T|^{(d-1)\nu}$, which for $d=3$ yields a value close to 1.25.

The interfacial profile $\rho(z)$ is obtained by integrating the differential equation (7.92). Again, exploiting the analogy with mechanics, it is more convenient to use energy conservation to obtain

$$\sqrt{\frac{C}{\kappa}}\mathrm{d}z = \frac{\mathrm{d}\rho}{(\rho^{(1)}-\rho)(\rho-\rho^{(2)})}. \tag{7.97}$$

It is convenient to change the variable to $\Delta\rho = \rho - (\rho^{(1)}+\rho^{(2)})/2$. One then integrates straighforwardly both sides, and obtains

$$\rho(z) = \frac{1}{2}\left(\rho^{(1)}+\rho^{(2)}\right) + \frac{1}{2}\left(\rho^{(2)}-\rho^{(1)}\right)\tanh\left(\frac{z}{2\xi}\right), \tag{7.98}$$

where the interfacial width ξ is given by

$$\xi = \frac{\sqrt{\kappa/C}}{|\rho^{(2)}-\rho^{(1)}|}. \tag{7.99}$$

This length diverges in approaching the critical point. This approach is valid in principle for ξ much larger than characteristic molecular sizes, but it turns out that its results are good enough also far from the critical point.

7.4 Virial Expansion

The interaction potential between particles exhibits a hard core, which makes a systematic expansion in which the interaction is treated as a perturbation impractical. If the particle density is low, however, it can play the role of a small parameter; one thus obtains a systematic expansion of the thermodynamic properties as a function of density, which is known as the **virial expansion**.

In this approach, the equation of state assumes the form

$$\frac{p}{k_B T} = \sum_{\ell=1}^{\infty} B_\ell(T)\, \rho^\ell, \tag{7.100}$$

where the coefficients $B_\ell(T)$ are called **virial coefficients**. Since the equation of state for ideal gases holds in the limit $\rho \to 0$, one obviously has

$$B_1(T) = 1, \qquad \forall T. \tag{7.101}$$

The expression of the virial coefficients as a function of T and the pair potential $u(r)$ is one of the first successes of "modern" statistical mechanics, and was derived by Mayer and Mayer [Maye40, chap. 13]. In this section, I will attempt to show, as a general outline, how to obtain it, and I will refer to the volume by Hansen and McDonald [Hans13, chap. 3] for further details.

Let us consider the canonical partition function Z_N:

$$Z_N = \frac{1}{N!} \frac{Q_N[v]}{\lambda_B^{3N}}. \tag{7.102}$$

If $v = v_{\mathcal{B}}$, given the expression of $Q_N[v]$, it can be written

$$Z_N = \frac{1}{N!} \int \prod_{i=1}^{N} \left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_B^3} \right) \prod_{(jk)} \left(\mathrm{e}^{-u\left(|\boldsymbol{r}_j - \boldsymbol{r}_k|\right)/k_B T} \right), \tag{7.103}$$

where the integral over each $\boldsymbol{r}_i$ is limited to the box $\mathcal{B}$. The last product runs over all different pairs (jk) of particles. Now, each factor $\mathrm{e}^{-u(r)/k_B T}$ is approximately equal to 1 as soon as $u(r) \ll k_B T$, which occurs as soon as $r \gg r_0$, where r_0 is the radius of the hard-core repulsion. For each pair (jk) of particles, we can therefore introduce the Mayer function f_{jk}, defined by

$$f_{jk} = \mathrm{e}^{-u\left(|\boldsymbol{r}_j - \boldsymbol{r}_k|\right)/k_B T} - 1. \tag{7.104}$$

It is easy to see that this function is bounded, equal to -1 at the origin, and goes to zero very rapidly when $r \gg r_0$. A plot of the Mayer function for the Lennard-Jones potential (and for $k_B T = \epsilon_0$) is shown in figure 7.2. The product in eq. (7.103) can therefore be written

$$\prod_{(jk)} (1 + f_{jk}) = 1 + \sum_{(jk)} f_{jk} + \sum_{(jk)} \sum_{(\ell m)} f_{jk} f_{\ell m} + \cdots. \tag{7.105}$$

The second term of this expression vanishes as soon as $r_{jk} = \left|\boldsymbol{r}_j - \boldsymbol{r}_k\right| \gg r_0$. Analogously, the third term vanishes if $r_{jk} \gg r_0$ or $r_{\ell m} \gg r_0$. Let us note that we can distinguish two types

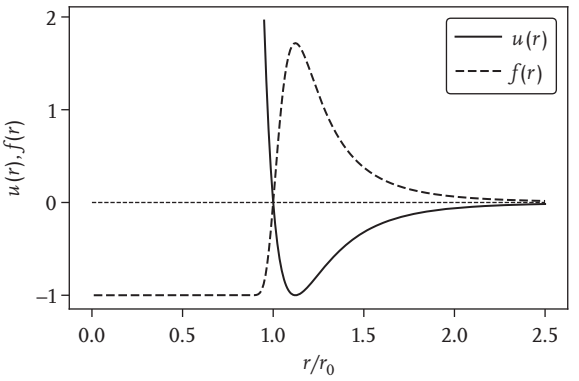


FIGURE 7.2. The Lennard-Jones potential $u(r)$ defined in eq. (7.11) and its Mayer function $f(r)$, evaluated for $k_B T = \epsilon_0$, as a function of r/r_0.

of contributions in the third term: those that concern four distinct particles, and those in which, for example, $\ell = j$, and which therefore concern only three particles.

If we stop at the first contribution, we obtain the following approximation for Z_N:

$$Z_N = \frac{1}{N!}\left(\frac{V}{\lambda_B^3}\right)^N \left[1 + \frac{N(N-1)}{V}\int \mathrm{d}\boldsymbol{r} f(r)\right]. \tag{7.106}$$

The Helmholtz free energy is given by $F = -k_B T \ln Z_N$. We thus obtain

$$\begin{aligned} F &= -k_B T \ln Z_N = F_0 - k_B T \ln\left[1 + \frac{N(N-1)}{V}\int \mathrm{d}\boldsymbol{r} f(r) + \cdots\right] \\ &\simeq F_0 - k_B T \frac{N^2}{2V}\int \mathrm{d}\boldsymbol{r} f(r), \end{aligned} \tag{7.107}$$

where $F_0 = -Nk_B T \ln(\mathrm{e}V/N\lambda_B^3)$ is the free energy of the ideal gas. By taking the derivative with respect to V, we obtain the pressure:

$$p = -\left.\frac{\partial F}{\partial V}\right)_{T,N} = k_B T\left[\rho - \frac{\rho^2}{2}\int \mathrm{d}\boldsymbol{r} f(r)\right]. \tag{7.108}$$

In this fashion, we obtain the expression of the second virial coefficient:

$$B_2(T) = -\frac{1}{2}\int \mathrm{d}\boldsymbol{r} f(r). \tag{7.109}$$

Exercise 7.4 By comparing (7.108) with (7.24), show that in this approximation $g(r)$ is expressed by

$$g(r) = 1 + f(r).$$

We now examine the behavior of $B_2(T)$. At high temperatures, in other words for $\epsilon_0 \ll k_B T$, we can approximate the Mayer function for a potential with a hard core as follows:

$$f(r) = \begin{cases} -1, & \text{for } r < r_0, \\ -u(r)/k_B T, & \text{for } r > r_0. \end{cases} \tag{7.110}$$

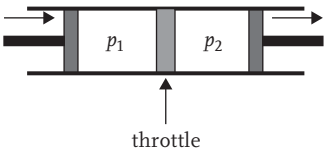


FIGURE 7.3. The Joule-Thomson process. The gas is pushed from the region at pressure p_1 to the region at pressure $p_2 < p_1$ through the throttle.

Therefore,

$$B_2(T) = -\frac{1}{2}\int \mathrm{d}\boldsymbol{r}\, f(r) \simeq b - \frac{a}{k_\mathrm{B}T}, \tag{7.111}$$

where

$$b = \frac{1}{2}\frac{4\pi}{3} r_0^3, \tag{7.112}$$

$$a = -\frac{1}{2}\int_{r_0}^{\infty} \mathrm{d}\boldsymbol{r}\, u(r) = -2\pi \int_{r_0}^{\infty} \mathrm{d}r\, r^2\, u(r). \tag{7.113}$$

These constants are positive and can be compared to those that appear in the van der Waals equation. More particularly, we see that, since r_0 is equal to twice the radius of the hard core of a gas particle, b is equal to four times its volume.

We also see that $B_2(T)$ is positive at high temperatures but decreases as T decreases. One exploits this fact in a gas cooling system, based on the **Joule-Thomson process**.

Exercise 7.5 (Joule-Thomson process) Consider a gas that expands through a throttle from a region at pressure p_1 to a region at pressure p_2, as shown in figure 7.3.

1. Show that the process occurs at constant enthalpy H.
2. Evaluate the temperature variation at constant enthalpy $\partial T/\partial p)_H$, as a function of both the specific heat at constant pressure C_p and the expansion coefficient $\partial V/\partial T)_p$ for a small pressure drop.
3. Given the equation of state expressing p as a power series in ρ truncated at ρ^2, express this quantity as a function of the second virial coefficient $B_2(T)$, and evaluate the temperature T^* in which $\partial T/\partial p)_H$ changes sign.
4. Evaluate the relation between the inversion temperature T^* and the pressure p for a van der Waals gas.

The second virial coefficient for the Lennard-Jones potential can be expressed in terms of a series expansion. We have, indeed,

$$\begin{aligned} B_2(T) &= -\frac{1}{2}\int \mathrm{d}\boldsymbol{r}\, f(\boldsymbol{r}) \\ &= -\frac{1}{2}\int_0^{\infty} \mathrm{d}r\, 4\pi r^2 \left\{ \exp\left[-\frac{4\epsilon_0}{k_\mathrm{B}T}\left(\left(\frac{r_0}{r}\right)^{12} - \left(\frac{r_0}{r}\right)^{6} \right) \right] - 1 \right\}. \end{aligned} \tag{7.114}$$

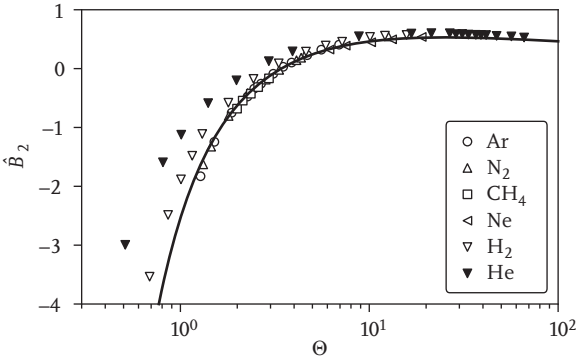


FIGURE 7.4. The reduced second virial coefficient $\hat{B}_2(\Theta)$ for the Lennard-Jones gas, as a function of the reduced temperature Θ, evaluated from eq. (7.117) (full curve), along with the experimental values for a number of gases. After [Hirs54].

Introducing b via eq. (7.112), and the notations

$$x = \frac{r}{r_0}, \qquad \Theta = \frac{k_B T}{\epsilon_0}, \qquad \widehat{B}_2(T) = \frac{B_2(T)}{b}, \tag{7.115}$$

we obtain, integrating by parts,

$$\begin{aligned} \widehat{B}_2(\Theta) &= \frac{4}{\Theta} \int_0^\infty \mathrm{d}x\, x^2 \left(\frac{12}{x^{12}} - \frac{6}{x^6} \right) \exp\left[-\frac{4}{\Theta} \left(x^{-12} - x^{-6} \right) \right] \\ &= \frac{4}{\Theta} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{4}{\Theta} \right)^n \int_0^\infty \mathrm{d}x\, x^2 \left(\frac{12}{x^{12}} - \frac{6}{x^6} \right) \exp\left(-\frac{4}{\Theta\, x^{12}} \right) x^{-6n}. \end{aligned} \tag{7.116}$$

The integrals can be expressed in terms of the Euler gamma function. By regrouping terms with the same power of Θ, one has [Hirs54, eq. (6.4)]

$$\widehat{B}_2(\Theta) = -\sum_{n=0}^{\infty} \frac{2^{n+1/2}}{4\, n!} \Gamma_E \left(\frac{2n-1}{4} \right) \Theta^{-(2n+1)/4}. \tag{7.117}$$

The coefficient approaches 0 (rather than 1) for $\Theta \to \infty$, since the Lennard-Jones potential does not, strictly speaking, have a hard core. This expression agrees satisfactorily with data on rare and some molecular gases, as shown in fig. 7.4, except for hydrogen and helium, due to quantum effects.

7.4.1 Higher Virial Coefficients

In order to evaluate further virial coefficients, it is convenient to use the grand canonical ensemble. Let us write the grand canonical partition function in the form (7.7):

$$Z_{GC} = \sum_{N=0}^{\infty} z^N Z_N = \sum_{N=0}^{\infty} \frac{z^N}{N!} Q_N, \tag{7.118}$$

where $z = e^{\mu/k_B T}$ is the fugacity. By substituting the expansion (7.105) in the expression (7.103) of Z_N, we obtain a sum of terms that can be expressed by diagrams. A diagram with N particles contains N vertices, labeled $i = 1, \ldots, N$, and a certain number of lines (ij) that connect vertex i with vertex j. There can be no more than one line that directly connects two given vertices. The contribution of a diagram of this type, with N vertices, is equal to

$$\frac{z^N}{N!} \int_{\mathcal{B}} \prod_{i=1}^{N} \left(\frac{d\boldsymbol{r}_i}{\lambda_B^3} \right) \prod_{(ij)} f_{ij},$$

where the product runs over all lines (ij) present in the diagram, and the integrals extend over the box $\mathcal{B}$. Evidently, each vertex j that is not connected by any line contributes a factor V/λ_B^3. If we therefore consider the contributions of the diagrams that contain a total of N vertices, but only k of which are connected by one or more lines, the corresponding contribution is equal to

$$\begin{aligned} &\frac{z^N}{N!} \frac{N!}{k!(N-k)!} \left(\frac{V}{\lambda_B^3} \right)^{N-k} \int_{\mathcal{B}} \prod_{i=1}^{k} \left(\frac{d\boldsymbol{r}_i}{\lambda_B^3} \right) \prod_{(ij)} f_{ij} \\ &\qquad = \frac{z^{N-k}}{(N-k)!} \left(\frac{V}{\lambda_B^3} \right)^{N-k} \left(\frac{z^k}{k!} \right) \int_{\mathcal{B}} \prod_{i=1}^{k} \left(\frac{d\boldsymbol{r}_i}{\lambda_B^3} \right) \prod_{(ij)} f_{ij}. \end{aligned}$$

The combinatorial factor $N!/(k!(N-k)!$ is a result of the different ways in which we can choose the k particles that are connected from among the N that are available. By summing over N, for a predetermined diagram, the first factor reproduces the Z_{GC}^0. We thus obtain

$$Z_{GC} = Z_{GC}^0 \left[1 + \sum_k \frac{z^k}{k!} \sum_{\mathcal{G}(k)} \int_{\mathcal{B}} \prod_{i \in \mathcal{G}(k)} \left(\frac{d\boldsymbol{r}_i}{\lambda_B^3} \right) \prod_{(ij) \in \mathcal{G}(k)} f_{ij} \right] = Z_{GC}^0 Z'_{GC}, \tag{7.119}$$

where $\mathcal{G}(k)$ runs over all the diagrams with k vertices, the product over i runs over all the vertices in the diagram $\mathcal{G}(k)$, and that over (ij) runs through all the lines that appear in the diagram. Let us note that the factor $1/k!$ is removed by summing over all the ways in which the k vertices can be labeled, and we can therefore remove it if we only consider diagrams in which the vertices are not labeled. In certain cases, even if the labels of certain particles are exchanged, the connections that appear in the diagram remain the same. The simplest case is that in which only the line (ij) appears, which clearly remains invariant with respect to the exchange $i \leftrightarrow j$. In these cases, in order to avoid multiple counts, it is necessary to divide the diagram's contribution by the number S of exchanges that leave the diagram unchanged. In order to evaluate S, it is necessary to label the diagram once. Once S has been obtained, one can forget about the labeling.

Let us now consider $\partial \ln Z_{GC}/\partial z$. On the one hand, since $z = e^{\mu/k_B T}$, we have

$$\left. \frac{\partial \ln Z_{GC}}{\partial z} \right)_T = \frac{1}{z} \left. \frac{\partial \ln Z_{GC}}{\partial \ln z} \right)_T = \frac{1}{z} N. \tag{7.120}$$

On the other hand, by taking the derivative of both sides of eq. (7.119), we obtain

$$\frac{\partial \ln Z_{\mathrm{GC}}}{\partial z}\Big)_T = \frac{V}{\lambda_{\mathrm{B}}^3} + \frac{1}{Z'_{\mathrm{GC}}} \sum_k \frac{z^k}{k!} \sum_{\mathcal{G}^0(k)} \int_{\mathcal{B}} \prod_{i \in \mathcal{G}^0(k)} \left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_{\mathrm{B}}^3} \right) \prod_{(ij) \in \mathcal{G}^0(k)} f_{ij}, \tag{7.121}$$

where the sum runs over the diagrams $\mathcal{G}^0(k)$ with $k+1$ vertices, of which one—let us refer to it as 0—is marked: it corresponds to the vertex with respect to which we took the derivative. We call this vertex the **root** and denote it by means of a small, open circle. The term V/λ_{B}^3 comes from the derivative of $\ln Z_{\mathrm{GC}}^0$. Each one of the diagrams we are considering includes a certain number of vertices that are directly or indirectly connected to 0 by means of lines and possibly some other nonconnected subdiagrams. By focusing our attention on a given subdiagram connected to 0, with k vertices, and by summing over all the others, we obtain the product of that subdiagram by Z'_{GC}, which cancels out with the first factor. We thus obtain

$$\frac{\partial \ln Z_{\mathrm{GC}}}{\partial z}\Big)_T = \frac{V}{\lambda_{\mathrm{B}}^3} + \sum_k \frac{z^k}{k!} \sum_{\mathcal{G}^0(k)} \int_{\mathcal{B}} \prod_{i \in \mathcal{G}^0(k)} \left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_{\mathrm{B}}^3} \right) \prod_{(ij) \in \mathcal{G}^0(k)} f_{ij}, \tag{7.122}$$

where the sum runs over all the diagrams *connected* to the root 0. By introducing the notation

$$\frac{V}{\lambda_{\mathrm{B}}^3} b_k = \sum_{\mathcal{G}^0(k)} \int_{\mathcal{B}} \prod_{i \in \mathcal{G}^0(k)} \left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_{\mathrm{B}}^3} \right) \prod_{(ij) \in \mathcal{G}^0(k)} f_{ij}, \tag{7.123}$$

and taking eq. (7.120), we obtain

$$\frac{N}{z} = \frac{V}{\lambda_{\mathrm{B}}^3} \left[1 + \sum_k \frac{z^k}{k!} b_k \right]. \tag{7.124}$$

Therefore

$$\rho = \frac{N}{V} = \frac{1}{\lambda_{\mathrm{B}}^3} \left[z + \sum_k \frac{z^{k+1}}{k!} b_k \right]. \tag{7.125}$$

We have thus obtained the expansion of density in powers of the fugacity. In order to reach our goal—the expression of pressure as a function of the density—we must invert this relation and obtain the expression of pressure. As we will see, this operation has a clear meaning in terms of diagrams.

In a diagram, a vertex that by our removing it gives disconnected subdiagrams (see fig. 7.5) is called a **node** or an **articulation point**. The relative contributions to subdiagrams that are connected by a node factorize. Let us, for example, consider two subdiagrams $\mathcal{G}_1$ and $\mathcal{G}_2$, connected by a vertex in j_0. In the diagram's contribution, the lines $(i_1 j_1)$ that connect the vertices belonging to $\mathcal{G}_1$ to one another and to j_0, and the analogous lines $(i_2 j_2)$ that connect the vertices belonging to $\mathcal{G}_2$, appear, but the lines that connect the subdiagrams to each other do not. By setting $\boldsymbol{r}_{i_1}^0 = \boldsymbol{r}_{i_1} - \boldsymbol{r}_j$ for $i_i \in \mathcal{G}_1$, and analogously for $\mathcal{G}_2$, we see that we can freely integrate over the $\boldsymbol{r}_{i_1}$ and $\boldsymbol{r}_{i_2}$. The contribution of subdiagrams $\mathcal{G}_1$ and $\mathcal{G}_2$ factorizes.

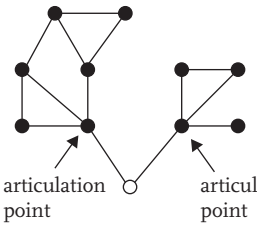


FIGURE 7.5. A diagram connected by two articulation points. The root is represented by an open circle.

Let us now consider a diagram with $k+1$ vertices in the expansion (7.122), in which there are no articulation points, except possibly the root. We can then obtain a set of diagrams with articulation points from this diagram, by "decorating" each of the k vertices with subdiagrams, except for the root. By summing with respect to all the "decorations," the contribution of the diagram gets multiplied by $(\rho\lambda_{\mathrm{B}}^3)^k$. In fact, the contribution of the decorations (with their respective z factors) reproduces N/z, up to the factor V/λ_{B}^3 that would correspond to the integration over the coordinates of the root, and the z factor is provided by the vertex to which the decorations are attached. We have therefore arrived at

$$\frac{N}{z}=\frac{V}{\lambda_{\mathrm{B}}^3}+\sum_k \frac{(\rho\lambda_{\mathrm{B}}^3)^k}{k!}\beta_k', \tag{7.126}$$

where

$$\beta_k'=\sum_{\hat{\mathcal{G}}(k)}\int_{\mathcal{B}}\prod_{i=\hat{\mathcal{G}}(k)}\left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_{\mathrm{B}}^3}\right)\prod_{(ij)\in\hat{\mathcal{G}}(k)} f_{ij} \tag{7.127}$$

is the contribution of all rooted diagrams of $k+1$ particles in which possibly only the root is an articulation point.

The structure of these diagrams is, generally, that of ℓ **irreducible** diagrams (in other words, diagrams free of articulation points), which join at the root. We will refer to the irreducible diagrams with $k+1$ vertices (of which one is the root) as $\mathcal{G}'(k)$, and to diagrams composed of ℓ irreducible diagrams—containing $k+1$ particles—which join at the root, as $\hat{\mathcal{G}}^{(\ell)}(k)$. One then has

$$\begin{aligned}\sum_k \frac{1}{k!}\sum_{\hat{\mathcal{G}}^{(\ell)}(k)}(\rho\lambda_{\mathrm{B}}^3)^k\int_{\mathcal{B}}\prod_{i\in\mathcal{G}'(k)}\left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_{\mathrm{B}}^3}\right)\prod_{(ij)\in\hat{\mathcal{G}}(k)} f_{ij}\\ =\frac{1}{\ell!}\frac{V}{\lambda_{\mathrm{B}}^3}\left[\sum_{\mathcal{G}(k)}(\rho\lambda_{\mathrm{B}}^3)^k\int_{\mathcal{B}}\prod_{0<i\in\mathcal{G}'(k)}\left(\prod\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_{\mathrm{B}}^3}\right)\prod_{(ij)\in\mathcal{G}'(k)} f_{ij}\right]^\ell.\end{aligned} \tag{7.128}$$

The factor $1/\ell!$ corrects the overcounting due to the exchange of irreducible diagrams with each other. Let us note that we have integrated explicitly over $\boldsymbol{r}_0$, thus obtaining the V/λ_{B}^3 factor. We now define β_k by

$$\beta_k=\sum_{\mathcal{G}'(k)}\frac{1}{k!}\int_{\mathcal{B}}\prod_{i\in\mathcal{G}'(k)}\left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_{\mathrm{B}}^3}\right)\prod_{(ij)\in\mathcal{G}'(k)} f_{ij}. \tag{7.129}$$

Using this definition, and taking the trivial term into account, we finally obtain

$$\frac{N}{z} = \frac{V}{\lambda_B^3} \exp\left[\sum_k \beta_k \left(\rho\lambda_B^3\right)^k\right]. \tag{7.130}$$

We have thus expressed the fugacity as a function of density. Let us take the logarithm of this expression:

$$\ln(\rho\lambda_B^3) - \frac{\mu}{k_B T} = \sum_k \beta_k \left(\rho\lambda_B^3\right)^k. \tag{7.131}$$

From the Gibbs-Duhem equation

$$\mathrm{d}\mu = -s\,\mathrm{d}T + \frac{1}{\rho}\,\mathrm{d}\rho, \tag{7.132}$$

we obtain

$$\left.\frac{\partial\mu}{\partial\rho}\right)_T = \frac{1}{\rho}\left.\frac{\partial p}{\partial\rho}\right)_T. \tag{7.133}$$

Therefore,

$$p(T,\rho) = \int_0^\rho \mathrm{d}\rho' \left.\frac{\partial p}{\partial\rho'}\right)_T = \int_0^\rho \mathrm{d}\rho'\,\rho \left.\frac{\partial\mu}{\partial\rho'}\right)_T. \tag{7.134}$$

But from eq. (7.131) we obtain

$$\left.\frac{\partial\mu}{\partial\rho'}\right)_T = \frac{k_B T}{\rho} - \lambda_B^3 k_B T \sum_k k\,\beta_k (\rho\lambda_B^3)^k. \tag{7.135}$$

By integrating, we finally obtain

$$p(T,\rho) = k_B T\,\rho - \lambda_B^3 k_B T \sum_k \frac{k}{k+1}\beta_k (\rho\lambda_B^3)^k. \tag{7.136}$$

This is the virial expansion of the equation of state. Evaluating the coefficients β_k is quite a challenge, since, as the order k increases, the number of diagrams increases more than exponentially. This expansion is fairly useful for determining the interaction potential starting from the equation of state, but it has scarce predictive capacity for liquids. In this case, it is important to find extrapolation schemes that somehow account for all the orders in the expansion of the density function. On the other hand, one would like to be able to evaluate more microscopic properties like $g(r)$, and not only the equation of state, in the liquid phase.

It can be shown that if the interaction potential satisfies certain stability and regularity conditions, the virial expansion (7.100) converges for $\rho < R_C$, where R_C depends on the properties of the potential and on the temperature. This analysis is at a higher mathematical level than that of the present book, and is detailed in [Ruel69, chap. 4], where one can find directions to the original literature.

For hard spheres, where the temperature dependence is trivial, one introduces the ratio of the volume occupied by the particles to the total volume of the system, called the

packing ratio:

$$\eta = \frac{1}{6}\pi\sigma^3\rho, \tag{7.137}$$

where σ is the diameter of the particle. Then the virial expansion takes the form

$$\frac{p}{\rho\, k_B T} = 1 + \sum_k \mathcal{B}_k \eta^k, \tag{7.138}$$

where the coefficients $\mathcal{B}_k$ are given by

$$\mathcal{B}_k = -\left(\frac{6}{\pi\sigma^3}\right)\frac{k}{k+1}\beta_k. \tag{7.139}$$

The first terms in the expansion (7.138) are given by [Hans13, p. 93] as

$$\begin{aligned}\frac{p}{\rho\, k_B T} = 1 + 4\eta + 10\,\eta^2 + 18.365\,\eta^3 + 28.224\,\eta^4 + 39.82\,\eta^5 \\ + 53.34\,\eta^6 + 68.54\,\eta^7 + 85.8\,\eta^8 + 105.8\,\eta^9 + \cdots .\end{aligned} \tag{7.140}$$

The truncated sum agrees well with molecular dynamics simulations for values of η up to 0.5.

It is interesting to note that it is possible to account almost exactly for the results obtained for the first virial coefficients for hard spheres by using an extrapolation formula due to Carnahan and Starling [Carn69] that gives

$$\frac{p}{\rho\, k_B T} = \frac{1 + \eta + \eta^2 - \eta^3}{(1-\eta)^3}. \tag{7.141}$$

The predictions of this equation differ from the Monte Carlo results by less than 0.3% over almost the entire fluid range.

7.4.2 Integral Equations for the g(r)

In this section, we derive some approximate integral equations that provide relations between the pair distribution $g(r)$ and the pair potential $u(r)$.

We start from the definition of the $\rho^{(2)}$:

$$\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) \propto \int \mathrm{d}\boldsymbol{r}\, \exp\left[-\frac{U\left(\boldsymbol{r}^N\right)}{k_B T}\right] \delta(\boldsymbol{x}_1 - \boldsymbol{r}_1)\, \delta(\boldsymbol{x}_2 - \boldsymbol{r}_2). \tag{7.142}$$

We can isolate the contribution of the direct interaction between particles 1 and 2 in $U\left(\boldsymbol{r}^N\right)$:

$$\rho^{(2)}(\boldsymbol{x}_1, \boldsymbol{x}_2) \propto \exp\left(-\frac{u(r_{12})}{k_B T}\right) \int \prod_{i=3}^{N} \mathrm{d}\boldsymbol{r}_i\, \exp\left[-\frac{U'\left(\boldsymbol{r}^N\right)}{k_B T}\right] \delta(\boldsymbol{x}_1 - \boldsymbol{r}_1)\, \delta(\boldsymbol{x}_2 - \boldsymbol{r}_2), \tag{7.143}$$

where $U'\left(\boldsymbol{r}^N\right)$ contains the sum of $u(r_{ij})$ over all pairs of particles, except for 1 and 2. We can therefore represent $\rho^{(2)}$ by means of a sum over diagrams, analogous to those we used for

the virial expansion. By some additional considerations, one easily arrives at the following result:

$$\rho^2 g(r)\, \mathrm{e}^{u(r_{12})/k_\mathrm{B}T} = \left(\frac{V}{\lambda_\mathrm{B}^3}\right)^2 \left[1 + \sum_{N=1}^{\infty} \frac{z^N}{N!} \int \prod_{i=1}^{N} \left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_\mathrm{B}^3}\right) \prod_{(ij)}{}' \left(1 + f_{ij}\right)\right], \tag{7.144}$$

where $r = r_{12} = |\boldsymbol{x}_1 - \boldsymbol{x}_2|$, and where the product runs over all the pairs of particles, except for (12). We can associate the usual diagrams with this expression, in which there are, however, now *two* marked points, namely 1 and 2, over which we do not integrate. A direct line between 1 and 2, moreover, is not allowed. By summing all the decorations of vertices 1 and 2, and keeping everything else as it is, we obtain the ρ^2 factor, which is removed by passing from the $\rho^{(2)}$ to the $g(r)$. Therefore, $g(r)\, \mathrm{e}^{u(r)/k_\mathrm{B}T}$ is represented in terms of the diagrams containing the two marked points 1 and 2, without the decorations of vertices 1 and 2, and without the direct line between these points. In addition to the trivial diagram composed of points 1 and 2 in isolation, all the other diagrams connect these points. Let us attempt to classify these diagrams.

We call a **parallel connection** a diagram that can be separated into subdiagrams that have the following properties:

1. Each subdiagram connects 1 and 2.
2. Different subdiagrams are not connected to one another, except in 1 or 2.

A diagram that does not contain parallel connections is called an **irreducible channel**. So more generally, we will be able to write

$$g(r)\, \mathrm{e}^{u(r_{12})/k_\mathrm{B}T} = 1 + \sum_k \frac{1}{k!} \left(\sum \text{irreducible channels}\right)^k. \tag{7.145}$$

The factor $1/k!$ comes, as usual, from the necessary correction to avoid multiple counting. Therefore,

$$\ln g(r) + \frac{u(r)}{k_\mathrm{B}T} = \sum \text{irreducible channels}. \tag{7.146}$$

We can identify the articulation points in the irreducible channels, as we did earlier. If a specific vertex is an articulation point, we distinguish between the decorations of the vertex——in other words, the subdiagrams that are connected to the rest of the diagram only through that vertex—and the subdiagrams that connect it to 1 and 2. By eliminating the decorations, we shift from the expansion in powers of z to that in powers of ρ.

Let us now consider the irreducible channels, free of decorations, that connect 1 and 2. We can distinguish the diagrams into two classes: the diagrams that contain articulation points, called **series** diagrams, and those that do not contain them, called **bridge** diagrams. We refer to the sum of all series diagrams as $\mathcal{S}$ and to the sum of all bridge diagrams as $\mathcal{B}$. We therefore have

$$\ln g(r) + \frac{u(r)}{k_\mathrm{B}T} = \mathcal{S} + \mathcal{B}. \tag{7.147}$$

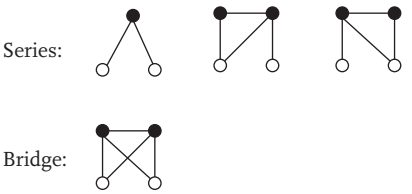


FIGURE 7.6. The first series and bridge diagrams. Series diagrams posses articulation points, whereas bridge diagrams do not.

The first diagrams that contribute to $\mathcal{S}$ and to $\mathcal{B}$ are shown in figure 7.6. Our strategy is to identify classes of diagrams that can be resummed so as to express $\mathcal{S}$ and $\mathcal{B}$ in terms of the $g(r)$ itself. In this fashion, one obtains closed equations for the $g(r)$ that can be solved numerically if necessary.

Let us consider the expansion of $h(r) = g(r) - 1$. It is easy to see that one has

$$h(r_{12}) = \sum \text{diagrams that connect 1 and 2, without decorations.} \tag{7.148}$$

Among these diagrams, we can identify diagrams with articulation points and those without articulation points. We refer to the sum of diagrams without articulation points as $c(r_{12})$. One then has

$$h(r_{12}) = c(r_{12}) + \rho \int \mathrm{d}\boldsymbol{r}_i \, c(r_{1i}) h(r_{i1}), \tag{7.149}$$

where i is the first articulation point one finds coming from 1. If we compare this relation with eq. (7.50), we see that the $c(r)$ that is so defined is the direct correlation function.

If we now assume that we can disregard the *bridge* diagrams, we see that the series diagrams reproduce only the second term of eq. (7.149): in other words, the $h(r) - c(r)$ term. We thus obtain the equation

$$\mathcal{S} \simeq h(r) - c(r), \tag{7.150}$$

which corresponds to

$$\ln g(r) + \frac{u(r)}{k_{\mathrm{B}} T} = h(r) - c(r). \tag{7.151}$$

Since $h(r)$ and $c(r)$ can be expressed as a function of $g(r)$, this is a closed integral equation in $g(r)$. It is known as the **hypernetted chain equation** (HNC) [Mori58].

In order to understand the meaning of this approximation, let us consider the first terms of the expansion of $c(r)$ in figure 7.7. We can write

$$g_{\mathrm{HNC}}(r) = \mathrm{e}^{-u(r)/k_{\mathrm{B}}T} \mathrm{e}^{\mathcal{S}} = (1 + f_{12}) \left(1 + \mathcal{S} + \frac{1}{2!} \mathcal{S}^2 + \cdots \right), \tag{7.152}$$

and therefore,

$$h_{\mathrm{HNC}}(r) = f_{12} \left(1 + \mathcal{S} + \frac{1}{2!} \mathcal{S}^2 + \cdots \right) + \mathcal{S} + \frac{1}{2!} \mathcal{S}^2 + \cdots, \tag{7.153}$$

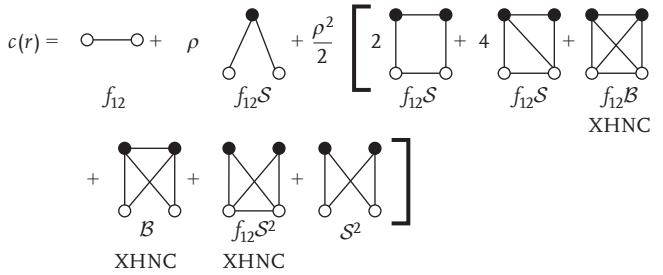


FIGURE 7.7. The first few terms of the expansion of $c(r)$ in powers of ρ. The diagrams labeled XHNC are neglected in the HNC approximation.

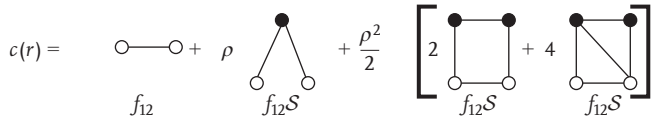


FIGURE 7.8. The first diagrams that contribute to the Percus-Yevick equation in an expansion in powers of ρ.

from which we obtain

$$h_{\mathrm{HNC}}(r) = f_{12}\left(1 + \mathcal{S} + \frac{1}{2!}\mathcal{S}^2 + \cdots\right) + \frac{1}{2!}\mathcal{S}^2 + \cdots . \tag{7.154}$$

A more drastic approximation (but, strangely, a more effective one) is obtained by neglecting the terms that contain powers of $\mathcal{S}$ greater than 1 in this last equation. One thus obtains

$$c_{\mathrm{PY}}(r) = f_{12}(1 + \mathcal{S}). \tag{7.155}$$

From this, it follows that

$$h_{\mathrm{PY}}(r) = c_{\mathrm{PY}}(r) + \mathcal{S} = f_{12} + \mathcal{S} + f_{12}\mathcal{S}, \tag{7.156}$$

and therefore,

$$\begin{aligned} g_{\mathrm{PY}}(r) &= 1 + f_{12} + \mathcal{S} + f_{12}\mathcal{S} = (1 + f_{12})(1 + \mathcal{S}) \\ &= \mathrm{e}^{-u(r)/k_{\mathrm{B}}T}\,[g_{\mathrm{PY}}(r) - c_{\mathrm{PY}}(r)]. \end{aligned} \tag{7.157}$$

This equation is known as the **Percus-Yevick equation** [Perc58], and is analytically solvable for pair potentials made of a hard core plus a short-range attractive tail [Wert64]. In figure 7.8, the first diagrams that contribute to the Percus-Yevick equation are shown. The results of the various approximations for the hard-sphere equation of state are plotted in figure 7.9, whereas fig. 7.10 shows the structure factor $S(k)$ of the hard-sphere gas as a function of $k\sigma$ for different values of the packing ratio $\eta = \pi\rho\sigma^3/6$, where σ is the diameter of the sphere obtained from the exact solution of the Percus-Yevick approximation for hard spheres.

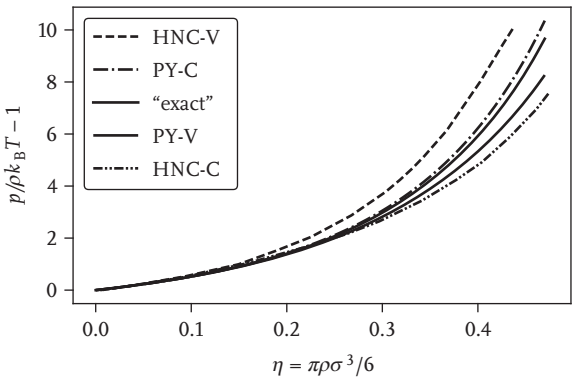


FIGURE 7.9. Equation of state of a hard sphere gas, obtained by means of the hypernetted chain (HNC) approximation and the Percus-Yervick (PY) equation. The deviation of the pressure from its ideal-gas value, namely the quantity $p/(\rho\, k_B T) - 1$, is plotted as a function of the packing ratio $\eta = \pi\rho\sigma^3/6$, where σ is the diameter of the sphere. The full line, with the "exact" designation, actually corresponds to the Carnahan-Starling extrapolation formula. The letters C and V refer to the two methods that allow one to obtain the equation of state from the $g(r)$: C refers to the compressibility equation (see exercise 7.2), whereas V refers to the pressure equation (7.24). After [Hans 13].

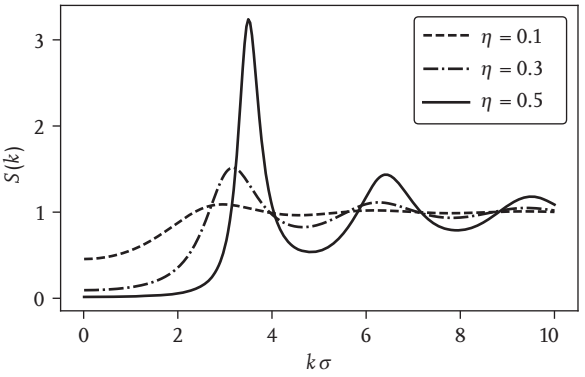


FIGURE 7.10. Structure factor $S(k)$ for the hard-sphere gas, obtained by means of the exact solution of the Percus-Yevick (PY) equation for hard spheres, due to M. S. Wertheim (1964). $S(k)$ is plotted as a function of $k\sigma$ for different values of the packing ratio $\eta = \pi\rho\sigma^3/6$, where σ is the diameter of the sphere.

7.5 Perturbation Theory

Another method to evaluate the behavior of liquids is based on an observation by Zwanzig [Zwan54]—namely, that the properties of a liquid are, as a first approximation, determined by the hard core, and only slightly modified by the attractive tail of the interaction. One can therefore attempt to express the properties of a fluid by splitting the interaction potential $u(r)$ into a *known* part u_0 and a *perturbation* u_1, which is assumed to be small, and then calculating the deviations from the behavior with u_0 alone in powers of the perturbation. It is clear that the results we obtained so far suggest taking a hard-sphere potential as u_0, so as to be able to exploit the exceptional efficacy of the Carnahan-Starling formula (7.141). Given that the hard-sphere potential is infinite, however, this means that the u_1 perturbation is not small in general.

Barker and Henderson [Bark76] have shown that one can solve this problem by choosing a temperature-dependent hard-sphere diameter σ. Their prescription is

$$\sigma = \int_0^{r_0} \mathrm{d}r \left[1 - \mathrm{e}^{-u_0(r)/k_\mathrm{B}T}\right], \tag{7.158}$$

where this time $u_0(r)$ is the repulsive part of the Lennard-Jones potential, that is,

$$u_0(r) = \theta_\mathrm{H}(r_0 - r)\, u(r), \tag{7.159}$$

in which $\theta_\mathrm{H}(x)$ is the Heaviside step function. Practically speaking, a system of soft spheres with potential $u_0(r)$ behaves like a system of hard spheres with a diameter σ given by eq. (7.158).

We can now set up the perturbation theory calculations. Let us consider the canonical partition function

$$Z_N = \frac{1}{N!} \int \prod_{i=1}^{N} \left(\frac{\mathrm{d}\boldsymbol{r}_i}{\lambda_\mathrm{B}^3} \right) \exp\left\{ -\frac{1}{k_\mathrm{B}T} \sum_{(ij)} \left[u_0(r_{ij}) + \lambda u_1(r_{ij}) \right] \right\}, \tag{7.160}$$

where we have introduced a perturbation parameter λ, which will be set to 1 at the end of the calculations.

We expand $\ln Z_N$ in powers of λ:

$$\ln Z_N = \ln Z_N(\lambda = 0) + \lambda \left. \frac{\partial \ln Z_N}{\partial \lambda} \right|_{\lambda=0} + \mathrm{O}(\lambda^2). \tag{7.161}$$

We have

$$\left. \frac{\partial \ln Z_N}{\partial \lambda} \right|_{\lambda=0} = -\frac{N(N-1)}{2V^2 k_\mathrm{B}T} (\lambda_\mathrm{B}^3)^2 \langle u_1 \rangle_0 , \tag{7.162}$$

where $\langle \ldots \rangle_0$ is the average over the unperturbed system ($\lambda = 0$). One has

$$\langle u_1 \rangle_0 = V \int \mathrm{d}\boldsymbol{r}\, u_1(r)\, g^{(0)}(r), \tag{7.163}$$

where $g^{(0)}(r)$ is the pair distribution of the unperturbed system, which can be expressed in terms of the $g(r)$ of the hard-sphere gas. From these expressions, we obtain, for example, an expression for the Helmholtz free energy, reading

$$\begin{aligned} F(T, V, N) &= -k_\mathrm{B}T \ln Z_N = -k_\mathrm{B}T \left[\ln Z_N(\lambda{=}0) + \lambda \left. \frac{\partial \ln Z_N}{\partial \lambda} \right|_{\lambda=0} + \cdots \right] \\ &= F_0(T, V, N) + N \frac{\rho}{2} \int \mathrm{d}\boldsymbol{r}\, u_1(r)\, g^{(0)}(r) + \cdots , \end{aligned} \tag{7.164}$$

where $F_0(T, V, N)$ is the Helmholtz free energy of the unperturbed system. The $g(r)$ of a Lennard-Jones fluid calculated for $\rho = 0.85\, r_0^{-3}$ and $k_\mathrm{B}T = 0.72\, \epsilon_0$ (close to the triple point) at the first order in perturbation theory, compared to the results of numerical simulations, is shown in figure 7.11. The agreement is fairly good and shows that the ideas behind the approximations are not too much off the mark.

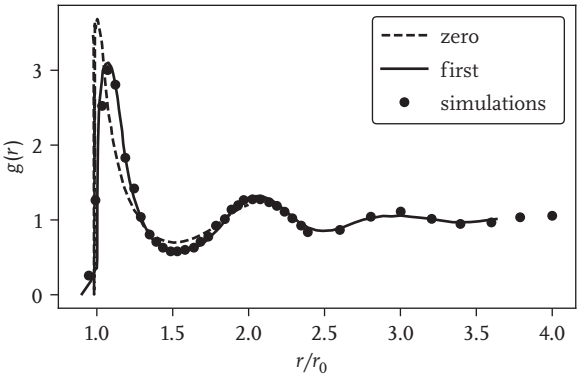


FIGURE 7.11. Pair distribution of a Lennard-Jones fluid for $\rho = 0.85\, r_0^{-3}$ and $k_B T = 0.72\, \epsilon_0$ (close to the triple point), at zero (dashed line) and first (solid line) orders in perturbation theory, compared with Verlet's numerical simulations. Adapted from [Bark76], with permission. (c) 1976 American Physical Society.

7.6 Liquid Solutions

Let us now consider a fluid composed of a prevalent chemical species (the **solvent**) in which a small quantity of another chemical species (the **solute**) is dissolved. If the concentration of the solute is fairly small, one is able to disregard the direct interactions between the molecules of the solute (but obviously not those between solute and solvent). Let us denote the number of particles of the solute and its chemical potential by n and μ, respectively, and those of the solvent by N and μ_s. The Gibbs free energy of the solution $G(T, p, N, n)$ contains a contribution proportional to n, as well as a contribution due to the entropy change, which derives from the fact that of the $(N+n)$ particles present in the solution, N are of one type and n of the other. This entropy change (called **entropy of mixing**) is given by

$$\begin{aligned} \Delta S &= k_B \ln \frac{(N+n)!}{N!n!} \simeq -k_B \left(N \ln \frac{N}{N+n} + n \ln \frac{n}{N+n} \right) \\ &\simeq -k_B \left(-n + n \ln \frac{n}{N} \right). \end{aligned} \tag{7.165}$$

Since G contains a term $-TS$, we obtain

$$G(T, p, N, n) = G(T, p, N, n=0) + n\, \psi(T, p, N) + n\, k_B T \ln \frac{n}{eN}, \tag{7.166}$$

where $\psi(T, p, N)$ is the enthalpy change per solute molecule. By taking the derivative of eq. (10.28) with respect to N and n, we obtain the corresponding chemical potentials:

$$\mu_s = \left. \frac{\partial G}{\partial N} \right)_{T,p,n} = \mu_s^0 - \chi\, k_B T, \tag{7.167}$$

$$\mu = \left. \frac{\partial G}{\partial n} \right)_{T,p,N} = \psi + k_B T \ln \chi, \tag{7.168}$$

where $\chi = n/(N+n) \simeq n/N$ is the molar fraction of the solute, and μ_s^0 is the chemical potential of the pure solvent.

Let us now consider the equilibrium of two regions of the solvent separated by a semipermeable membrane, which will allow the solvent, but not the solute, to pass. In this situation, the chemical potential of the solvent (but not of the solute) must be equal on both sides of the membrane. We thus obtain

$$\mu_s(T, p_1, \chi_1) = \mu_s^0(T, p_1) - k_B T\, \chi_1 = \mu_s^0(T, p_2) - k_B T\, \chi_2. \tag{7.169}$$

Expanding to the first order in the difference $\Delta p = p_2 - p_1$, we obtain

$$\mu_s^0(T, p_2) = \mu_s^0(T, p_1) + \Delta p \left.\frac{\partial \mu_s^0}{\partial p}\right)_T = \mu_s^0(T, p_1) + \Delta p\, v, \tag{7.170}$$

where $v = V/N$ is the volume per molecule of the pure solvent. (We have applied the Gibbs-Duhem equation.) Therefore,

$$\Delta p = k_B T\, \frac{\chi_2 - \chi_1}{v}. \tag{7.171}$$

More particularly, if on one of the two sides we have pure solvent, we obtain

$$\Delta p = k_B T\, \frac{\chi_1}{v} = k_B T\, \frac{n}{Nv} = k_B T\, \frac{n}{V}. \tag{7.172}$$

The osmotic pressure for a diluted solution therefore follows the same law as a perfect gas. This result is known as **van 't Hoff's law**.

7.6.1 Solvation Heat

The dissolving process takes place with absorption or release of heat. We wish to evaluate this thermal effect. Assume that the process takes place at constant temperature and pressure, and that the solute concentration is small. Let us thus consider a solution in the presence of the pure solute, and assume that a small number δn of the solute goes from the pure solute to the solution. Then the change δG of the Gibbs free energy of the whole system is given by the sum of its change in the solution and that in the pure solute. We thus have, for the solution,

$$\delta G_{sol} = \frac{\partial G_{sol}}{\partial n}\, \delta n = \mu\, \delta n, \tag{7.173}$$

where μ is the chemical potential of the solute in the solution. We have on the other hand, for the pure solute,

$$\delta G_0 = -\frac{\partial G_0}{\partial n}\, \delta n = -\mu_0\, \delta n, \tag{7.174}$$

where G_0 is the Gibbs free energy of the pure solute and μ_0 its corresponding chemical potential. We thus obtain the change in the total Gibbs free energy of the system:

$$\delta G = \delta n\, (\mu - \mu_0). \tag{7.175}$$

Substituting the expression (7.168) of μ in this equation, we obtain

$$\delta G = -\delta n\ (\psi - \mu_0 + k_B T \ln \chi) = -\delta n\, k_B T\ \ln \frac{\chi_0(p, T)}{\chi}, \tag{7.176}$$

where the **solubility** $\chi_0(p, T)$ is defined by

$$\chi_0(p, T) = \exp\left(\frac{\mu_0 - \psi}{k_B T}\right), \tag{7.177}$$

and is equal to the molar fraction of the saturated solution, namely, the one that is in equilibrium with the pure solute.

Thus $|\delta G|$ is the maximal work that one can obtain by dissolving δn particles of the pure solute, or, equivalently, the minimal work needed to purify δn solute particles.

We now evaluate the heat δQ absorbed during the process. This heat is equal to the enthalpy change δH. Now, by a straightforward differentiation, we have the following expression of H:

$$H = -T^2\ \frac{\partial}{\partial T}\frac{G}{T}\Big)_T. \tag{7.178}$$

We thus obtain

$$\delta H = -T^2\ \frac{\partial}{\partial T}\frac{\delta G}{T}\Big)_T = \delta n\, k_B T^2\ \ \frac{\partial \ln \chi_0}{\partial T}\Big)_p. \tag{7.179}$$

Integrating in δn (always assuming that the solution remains diluted) we obtain the **solvation heat** for a number n of solute molecules:

$$Q = \Delta H = \delta n\, k_B T^2\ \ \frac{\partial \ln \chi_0}{\partial T}\Big)_p. \tag{7.180}$$

We can also evaluate the volume change due to the dissolution process. We have

$$\delta V = \frac{\partial \delta G}{\partial p}\Big)_T. \tag{7.181}$$

Substituting (7.176), we obtain

$$\delta V = -\delta n\, k_B T\ \ \frac{\partial \ln \chi_0}{\partial p}\Big)_T. \tag{7.182}$$

Thus the absorbed heat is proportional to the change in solubility upon a temperature increase, and the volume change is proportional to the solubility change upon a pressure increase. Both expressions are in agreement with Le Châtelier's principle. Consider, for example, a saturated solution. If upon lowering the temperature, the solubility increases, some heat, $Q < 0$, since $\partial \chi_0 / \partial T < 0$ in this case will be released. The released heat acts as if opposing the cooling that perturbed the equilibrium of the system. A similar argument holds for eq. (7.182).

Finally, to evaluate ψ, we take advantage of the following method. Given the solution containing N particles of the solvent and n particles of the solute, we distinguish between the interaction potential $U\left(\boldsymbol{r}_s^N\right)$ between the solvent particles and the potential $U_{\text{mix}}\left(\boldsymbol{r}_s^N, \boldsymbol{r}^n\right)$

between the solvent and the solute particles. We do not need to take into account the interactions among the solute particles, because they happen so rarely. Here $\boldsymbol{r}_{\mathrm{s}}^{N} = (\boldsymbol{r}_{\mathrm{s},1}, \ldots, \boldsymbol{r}_{\mathrm{s},N})$ is the collection of the positions of the solvent particles, and $\boldsymbol{r}^{n} = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n)$ that of the solute particles.

We then introduce a parameter λ and evaluate the canonical partition function of a system of $(N+n)$ particles with an interacting potential given by $U(\boldsymbol{r}_{\mathrm{s}}, \boldsymbol{r}; \lambda) = U(\boldsymbol{r}_{\mathrm{s}}) + \lambda U_{\mathrm{mix}}(\boldsymbol{r}_{\mathrm{s}}, \boldsymbol{r})$. We obtain

$$\begin{aligned} Z_{N+n}(\lambda) &= Z_{N+n}^{(0)} \int \mathrm{d}\boldsymbol{r}_{\mathrm{s}}^{N} \int \mathrm{d}\boldsymbol{r}^{n} \, \exp\left[-\frac{1}{k_{\mathrm{B}}T}\left(U\left(\boldsymbol{r}_{\mathrm{s}}^{N}\right) + \lambda U_{\mathrm{mix}}\left(\boldsymbol{r}_{\mathrm{s}}^{N}, \boldsymbol{r}^{n}\right)\right)\right] \\ &= Z_{N+n}^{(0)} \, Q_{N+n}(\lambda), \end{aligned} \tag{7.183}$$

where $Z_{N+n}^{(0)}$ is the partition function of the ideal gas of $(N+n)$ particles, and $Q_{N+n}(\lambda)$ is the configuration integral. Let us now evaluate the change in $\ln Z$ as λ grows from 0 to 1. Since $F = -k_{\mathrm{B}}T \ln Z$, this is proportional to the reversible work needed to dissolve n solute particles in the system. We obtain

$$\begin{aligned} \frac{\partial \ln Z_{N+n}}{\partial \lambda} = \frac{1}{Q_{N+n}(\lambda)} \int \mathrm{d}\boldsymbol{r}_{\mathrm{s}}^{N} \int \mathrm{d}\boldsymbol{r}^{n} \left(-\frac{U_{\mathrm{mix}}\left(\boldsymbol{r}_{\mathrm{s}}^{N}, \boldsymbol{r}^{n}\right)}{k_{\mathrm{B}}T}\right) \\ \times \exp\left[-\frac{1}{k_{\mathrm{B}}T}\left(U\left(\boldsymbol{r}_{\mathrm{s}}^{N}\right) + \lambda U_1\left(\boldsymbol{r}_{\mathrm{s}}^{N}, \boldsymbol{r}^{n}\right)\right)\right]. \end{aligned} \tag{7.184}$$

We now express $U_{\mathrm{mix}}\left(\boldsymbol{r}_{\mathrm{s}}^{N}, \boldsymbol{r}^{n}\right)$ by

$$U_{\mathrm{mix}}\left(\boldsymbol{r}_{\mathrm{s}}^{N}, \boldsymbol{r}^{n}\right) = \sum_{i=1}^{N} \sum_{j=1}^{n} u_{\mathrm{mix}}(|\boldsymbol{r}_{\mathrm{s},i} - \boldsymbol{r}_j|). \tag{7.185}$$

Substituting in (7.184), we obtain, after some manipulations recalling those of sec. 7.4,

$$\frac{\partial \ln Z_{N+n}}{\partial \lambda} = -\frac{n}{k_{\mathrm{B}}T} \rho_{\mathrm{s}} \int \mathrm{d}\boldsymbol{r} \, u_{\mathrm{mix}}(r) \, g_{\mathrm{mix}}(r; \lambda), \tag{7.186}$$

where ρ_{s} is the solvent concentration and $g_{\mathrm{mix}}(r; \lambda)$ is the radial distribution function for a solute-solvent pair, for the given value of λ. Formally integrating the derivative, we obtain

$$\begin{aligned} F(V, T, N, n) = F^{(0)}(V, T, N, n) + \Delta F_{\mathrm{s}}(V, T, N) \\ + n\rho_{\mathrm{s}} \int_0^1 \mathrm{d}\lambda \int \mathrm{d}\boldsymbol{r} \, u_{\mathrm{mix}}(r) \, g_{\mathrm{mix}}(r; \lambda), \end{aligned} \tag{7.187}$$

where $F^{(0)}$ is the ideal gas free energy of a mixture of $(N+n)$ particles and ΔF_{s} is the excess free energy of the pure solvent due to its interactions. Taking the derivative with respect to n, we finally obtain eq. (7.168) in the form

$$\mu = \frac{\partial F^{(0)}}{\partial n} + \psi = k_{\mathrm{B}}T \, \ln \chi + \psi, \tag{7.188}$$

where ψ is given by

$$\psi = \rho_s \int_0^1 d\lambda \int d\boldsymbol{r}\, u_{mix}(r)\, g_{mix}(r;\lambda). \tag{7.189}$$

7.6.2 Ionic Solutions

Van 't Hoff's law is not valid if the interactions between the molecules of the solute are long range. In this case, the free energy exhibits singular terms in the concentration of the solute. In practice, the most important case is that of ionic solutions. In order to clarify our reasoning, let us assume that our salt is formed by a single cation and a single anion, where each one is monovalent, and that we denote the concentration of the cations (anions) by ν_+ (ν_-), both being equal, on average, to the molar concentration ν_0 of the salt. We now show that the effect of these free charges is to *screen* the Coulomb interactions in the solution beyond a characteristic length called the **Debye length**.

Let us now consider a charge q placed at the origin: it produces an electric field, represented by an electrostatic potential $\phi(\boldsymbol{r})$, that polarizes the solution. In order to evaluate this effect, let us consider that if the local charge density at a generic point $\boldsymbol{r}$ is equal to $\rho(\boldsymbol{r})$, then the potential $\phi(\boldsymbol{r})$ is a solution of the Poisson equation:

$$-\nabla^2\phi = \frac{\rho}{\epsilon_0\epsilon_r}. \tag{7.190}$$

On the other hand, if the electrostatic potential is equal to $\phi(\boldsymbol{r})$, the local density of ions with a $\pm e$ charge is proportional to the Boltzmann factor $e^{\mp e\phi(r)/k_B T}$. Since both the Poisson equation and the Boltzmann factor are present, this theory (which is due to Debye and Hückel [Deby23]) is also known as the **Poisson-Boltzmann theory**.

If the concentration ν is sufficiently small, to the extent that the $\phi(\boldsymbol{r})/k_B T$ factor can be considered small throughout the solution (except possibly for a microscopic volume around the point charges), we can expand the Boltzmann factor to the first order in $\phi(\boldsymbol{r})/k_B T$, and obtain

$$\nu_\pm(\boldsymbol{r}) = \nu_0 \left[1 \mp \frac{e\,\phi(\boldsymbol{r})}{k_B T}\right]. \tag{7.191}$$

But then the local charge density is equal to

$$\rho(\boldsymbol{r}) = e\,(\nu_+ - \nu_-) = 2\nu_0\, e^2\, \frac{\phi(\boldsymbol{r})}{k_B T}. \tag{7.192}$$

We thus obtain the equation for ϕ,

$$\nabla^2\phi = \frac{\phi}{\ell_D^2}, \tag{7.193}$$

where ℓ_D is the **Debye screen length**:

$$\ell_D^2 = \frac{\epsilon_0\epsilon_r\, k_B T}{2\nu_0\, e^2}. \tag{7.194}$$

The solution of this equation, with a charge q at the origin, is given by

$$\phi_q(r) = \frac{q}{4\pi\,\epsilon_0\epsilon_r}\,\frac{\mathrm{e}^{-r/\ell_D}}{r}. \tag{7.195}$$

We therefore see that Coulomb's law holds only for $r/\ell_D \ll 1$, after which the Coulomb interactions are screened. By substituting in eq. (7.192), we obtain the charge density $\rho(r)$ at a distance r from the origin:

$$\rho(r) = -2\nu_0\, e^2\,\frac{\phi(r)}{k_B T} = -\frac{q\nu_0 e^2}{2\pi\,\epsilon_0\epsilon_r}\,\frac{\mathrm{e}^{-r/\ell_D}}{r} = -\frac{q}{4\pi\,\ell_D^2}\,\frac{\mathrm{e}^{-r/\ell_D}}{r}. \tag{7.196}$$

By expanding the expression of the potential around an ion of charge q for small values of r/ℓ_D, we obtain

$$\phi(r) \simeq \frac{q}{4\pi\,\epsilon_0\epsilon_r}\left(\frac{1}{r} - \frac{1}{\ell_D}\right). \tag{7.197}$$

The first term represents the potential produced by the ion in question, while the second term represents the effect of its interactions with all the other ions in solution. The potential energy of the ion is therefore given by $-q^2/(4\pi\,\epsilon_0\epsilon_r\ell_D)$. The total electrostatic energy U_{el} is obtained by summing this expression over all the ions present in the solution. One thus obtains

$$U_{el} = -\frac{2V\nu_0\, e^2}{4\pi\,\epsilon_0\epsilon_r\ell_D} = -\frac{e^3 V}{4\pi}\left(\frac{2\nu_0}{\epsilon_0\epsilon_r}\right)^{3/2}(k_B T)^{-1/2}. \tag{7.198}$$

By integrating the thermodynamic relation

$$-\left.\frac{\partial(F/T)}{\partial T}\right)_{V,N} = \frac{F}{T^2} + \frac{S}{T} = \frac{E}{T^2}, \tag{7.199}$$

we obtain the correction that must be added to the Helmholtz free energy:

$$F_{el} = -\frac{e^3 V}{3\pi}\left(\frac{2\nu_0}{\epsilon_0\epsilon_r}\right)^{3/2}(k_B T)^{-1/2}. \tag{7.200}$$

We can now substitute ν_0 with its expression $\nu_0 = N_{salt}/V$. We thus obtain

$$F_{el} = -\frac{e^3}{3\pi}\left(\frac{2N_{salt}}{\epsilon_0\epsilon_r}\right)^{3/2}(V\,k_B T)^{-1/2}. \tag{7.201}$$

By taking the derivative of F with respect to V, we obtain the electrostatic pressure of the ions, which is proportional to $V^{-3/2}$ and to $T^{-1/2}$.

Exercise 7.6 Generalize this result to the case of a salt with composition $\sum_\alpha n_\alpha X_\alpha$, where the ion X_α has a charge $z_\alpha\, e$. (One obviously has $\sum_\alpha n_\alpha z_\alpha = 0$ due to neutrality.)

Let us note that the results we obtained also exhibit some paradoxical aspects if compared with the Ornstein-Zernike theory of $g(r)$. Because of the screening effect of the Coulomb potential, the correlation function $h(r)$ actually becomes *short range* (it vanishes over distances on the order of the Debye length), while both the pair potential and the direct correlation are long range.

7.7 Colloidal Suspensions

A number of interesting phenomena occur when **amphiphilic molecules** (also called **surfactants**) are dissolved in a solvent like water. These molecules are formed of a hydrophilic part and a hydrophobic part, bound together. As a consequence, they have the tendency to avoid remaining in the bulk of the solvent, and they gather either at the oil-water interface, with the hydrophilic head toward the water and the hydrophobic tail toward the oil, or at the water-air interface, with the hydrophobic tail hanging above the surface. When their concentration is sufficiently high, they also aggregate among themselves, forming aggregates of nano- or micrometer size, called **micelles**, where they can place their hydrophilic heads at the aggregate surfaces in contact with the solvent, and keep their hydrophobic tails in contact among themselves, away from the solvent. The resulting system is called a **colloidal suspension**. Such a suspension is usually **polydisperse**, that is, formed of aggregates containing different numbers of molecules. In the following, we consider the simple case in which only one species of amphiphilic molecules is present.

We first consider the effects of surfactants on the surface tension of the interface. Denote by A the area of the interface and by $\sigma_{\rm ad}$ the surface density of adsorbed surfactant molecules. Then the free energy of the interface is given by

$$F^{(\rm s)} = A\left\{\gamma_0 + k_{\rm B}T\sigma_{\rm ad}\left[\ln(\sigma_{\rm ad}\ell^2) - 1\right] + \sigma_{\rm ad}\,\epsilon_{\rm s}\right\}, \tag{7.202}$$

where γ_0 is the surface tension in the absence of surfactants, $\epsilon_{\rm s}$ is the energy per adsorbed surfactant molecule, and the second term in braces takes into account the ideal entropy of the adsorbed surfactant. Here ℓ is a length scale (akin to $\lambda_{\rm B}$) that plays no role in what follows. By taking the derivative with respect to A, at constant number $N_{\rm ad} = \sigma_{\rm ad}\,A$ of surfactant molecules, we obtain the surface tension:

$$\gamma = \frac{\partial F^{(\rm s)}}{\partial A}\bigg)_{T,N_{\rm ad}} = \gamma_0 - k_{\rm B}T\,\sigma_{\rm ad}. \tag{7.203}$$

Thus the surface tension is reduced by a quantity proportional to the surface density of the surfactants, which explains their name.

We now show that the aggregation into micelles takes place at a rather well-defined concentration, even if the phenomenon is not so sharp as to be described as a phase transition. Let us denote by N the total number of amphiphilic molecules, by N_α the total number of molecules belonging to micelles made of α molecules, and by $n_\alpha = N_\alpha/\alpha$ the number of micelles of size α. We obviously have

$$\sum_\alpha N_\alpha = N. \tag{7.204}$$

If the suspension is sufficiently diluted, we can neglect the interactions among aggregates. Then the partition function of the suspension is obtained by multiplying together the partition functions for the aggregates of each size α:

$$Z_N = \prod_\alpha \frac{1}{n_\alpha!}\left(\frac{V\,\zeta_\alpha^{\rm int}}{\lambda_{{\rm B}\alpha}^3}\right)^{n_\alpha}, \tag{7.205}$$

where $\zeta_\alpha^{\rm int}$ is the internal partition function of a micelle of size α, and $\lambda_{{\rm B}\alpha}$ its corresponding thermal de Broglie wavelength. Then the free energy of the suspension is given by

$$F = -k_{\rm B}T \ln Z_N = \sum_\alpha \left\{ n_\alpha \, k_{\rm B}T \ln \left[\left(\frac{n_\alpha \, \lambda_{{\rm B}\alpha}^{\,3}}{V} \right) - 1 \right] + n_\alpha \, f_\alpha^{\rm int} \right\}, \tag{7.206}$$

where $f_\alpha^{\rm int} = -k_{\rm B}T \ln \zeta_\alpha^{\rm int}$ is the internal free energy of a micelle of size α.

Thus, the chemical potential of an amphiphilic molecule belonging to a micelle of size α is given by

$$\mu_\alpha = \frac{\partial F}{\partial N_\alpha} \Big)_{T,V} = \frac{1}{\alpha} \, \frac{\partial F}{\partial n_\alpha} \Big)_{T,V} = \frac{k_{\rm B}T}{\alpha} \ln \left(\frac{n_\alpha \lambda_{{\rm B}\alpha}^{\,3}}{V} \right) + \frac{f_\alpha^{\rm int}}{\alpha}. \tag{7.207}$$

Denoting by $N_{\rm tot}$ the total number of molecules (amphiphilic and solvent) in the system, we can cast the expression of μ_α as follows:

$$\mu_\alpha = \frac{k_{\rm B}T}{\alpha} \ln \left(\frac{x_\alpha}{\alpha} \right) + \epsilon_\alpha, \tag{7.208}$$

where $x_\alpha = N_\alpha / N_{\rm tot}$ is the mole fraction of amphiphilic molecules that belong to micelles of size α, and

$$\epsilon_\alpha = \frac{1}{\alpha} \left[f_\alpha^{\rm int} - k_{\rm B}T \ln \left(\frac{v}{\lambda_{{\rm B}\alpha}^{\,3}} \right) \right], \tag{7.209}$$

where $v = V / N_{\rm tot}$ is the mean volume per molecule in the solution. For a diluted solution, such as the ones we consider here, v is practically equal to the volume per solvent molecule, and is therefore independent of the number of amphiphilic molecules in the system.

Chemical equilibrium for the amphiphilic molecules requires the chemical potential μ_α to have a value μ independent of α. We thus obtain

$$x_\alpha = \alpha \exp \left[\frac{\alpha}{k_{\rm B}T} \left(\mu - \epsilon_\alpha \right) \right]. \tag{7.210}$$

By exploiting the relation $x_1 = \exp \left[\left(\mu - \epsilon_1 \right) / k_{\rm B}T \right]$, this expression can be recast in the form

$$x_\alpha = \alpha x_1^\alpha \exp \left[\frac{\alpha}{k_{\rm B}T} \left(\epsilon_1 - \epsilon_\alpha \right) \right]. \tag{7.211}$$

The tendency of the surfactant to aggregate is expressed by the fact that $\epsilon_1 > \epsilon_\alpha$, at least for some values of α. For low values of x, most of the surfactant molecules remain isolated, and one has $x_1 \simeq x$. As x increases, so does μ, until it becomes close to the value of the smallest ϵ_α. From there on, most of the added surfactant molecules go into anggregates, and the values of x_1 and μ grow much more slowly. The value $x_{\rm c}$ of x where this takes place is called the **critical micellar concentration** (**cmc**). Conventionally, $x_{\rm c}$ is defined as the value of x where it is equally probable for an amphiphilic molecule to be a monomer or to belong in an aggregate:

$$\sum_{\alpha \geq 2} x_\alpha = x - x_1 = x_1. \tag{7.212}$$

The details of micellar organization are determined by the structure of the amphiphilic molecules. Molecules with a large hydrophylic head and a comparatively short hydrophobic tail have a roughly conical shape and tend to aggregate into more or less spherical micelles. In this case, ϵ_α has a pronounced minimum for a value α^*. If we assume that most of the aggregates have size α^*, the condition for x_c reads

$$\sum_{\alpha \geq} x_\alpha \simeq x_{\alpha^*} = x_1 = \frac{x_c}{2}. \tag{7.213}$$

Substituting in eq. (7.211) we obtain the relation

$$x_c = 2\,(\alpha^*)^{1/(\alpha^*-1)} \exp\left[\frac{\alpha^*}{(\alpha^*-1)k_B T}\,(\epsilon_{\alpha^*} - \epsilon_1)\right] \simeq 2\exp\left(\frac{\epsilon_{\alpha^*} - \epsilon_1}{k_B T}\right), \tag{7.214}$$

where we have exploited the fact that, usually, $\alpha^* \gg 1$. Typical values of x_c in this scenario are of order $10^{-5} \div 10^{-4}$, with typical aggregation numbers α^* of order 10^2. We can then express the relation between x_1 and x_c as follows:

$$x \simeq x_1 + \alpha^* \left(\frac{x_1}{x_c}\right)^{\alpha^*}. \tag{7.215}$$

Since $\alpha^* \gg 1$, a slight increase in x_1 produces a large increment in the second term of this equation. Therefore, as x increases, x_1 increases much more slowly, and remains close to x_c at least for accessible amphiphile concentrations.

A different scenario obtains when the molecules tend to aggregate in cylindrical micelles. In this case we expect ϵ_α to be of the form

$$\epsilon_\alpha = \epsilon_\infty + \frac{\delta f}{\alpha}, \tag{7.216}$$

where δf is the extra free energy associated with the roughly hemispherical caps needed to shield at both ends the interior of the cylinder from contact with the solvent. Substituting in eq. (7.211) we obtain

$$x_\alpha = \alpha \left[x_1 \exp\left(\frac{\delta f}{k_B T}\right)\right]^\alpha \exp\left(-\frac{\delta f}{k_B T}\right). \tag{7.217}$$

The normalization condition $\sum_{\alpha=1}^{\infty} x_\alpha = x$ yields

$$\sum_{\alpha=1}^{N} x_\alpha = \frac{x_1}{\left(1 - x_1 \exp(-\delta f / k_B T)\right)^2} = x. \tag{7.218}$$

Setting $x_c = 2x_1$ we obtain

$$x_c = \left(2 - \sqrt{2}\right) z, \tag{7.219}$$

where

$$z = \mathrm{e}^{-\delta f / k_B T}. \tag{7.220}$$

More generally, we obtain

$$x_1 = \frac{z^2}{2x}\left(1 + \frac{2x}{z} - \sqrt{1 + \frac{4x}{z}}\right). \tag{7.221}$$

For $x \ll z$ this yields $x_1 \simeq x$, whereas for $x \gg z$ it behaves as $z\,(1 - \sqrt{z/x})$, showing that the monomer concentration remains practically constant, as all excess amphiphilic molecules go into the micelles.

Recommended Reading

A concise introduction to the modern theory of classical and complex fluids is found in J.-L. Barrat and J.-P. Hansen, *Basic Concepts for Simple and Complex Liquids*, Cambridge University Press, Cambridge, UK, 2003, and a more extensive one in J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids with Applications to Soft Matter*, 4th ed., Amsterdam-New York: Academic Press, 2013. Intermolecular forces are treated in great detail in J. Israelachvili, *Intermolecular & Surface Forces*, 2nd ed., London: Academic Press, 1991. For an accessible introduction to soft matter, one can read M. Doi, *Soft Matter Physics*, Oxford University Press, Oxford, 2013. Liquid crystals are comprehensively treated in P.-G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, Oxford: Clarendon Press, 2nd ed., 1993. The physics of inhomogeneous fluids is treated, among others, by S. A. Safran, *Statistical Thermodynamics of Surfaces, Interfaces and Membranes*, Reading, MA: Addison-Wesley, 1994.

8 Numerical Simulation

Meditate, before thinking!

—Stanisław J. Lec

The invention of the computer has had a very significant impact on statistical mechanics. One the one hand, it has allowed us to tackle problems that require complicated calculations, and on the other, it has provided an important contribution to the understanding of fundamental concepts. In this chapter, we will briefly discuss some methods and problems related to the use of the computer in statistical mechanics.

8.1 Introduction

We must distinguish between two ways of using the computer in theoretical physics: **calculation** and **simulation**. In *calculation*, the computer performs manipulations that would be too long or tedious to be reliably performed manually. When faced with a method of perturbation theory, for instance, such as high-temperature expansions, it is possible to entrust the evaluation of higher order contributions to a computer. A classic example of the use of a computer for these purposes is Appel and Haken's proof of the four-color theorem [Appe77]. In this case, the theory predicted that the theorem would be proven if it was possible to find a certain class of configurations of regions in the plane that possessed a certain set of properties. If a specific class of configurations did not possess some of these properties, the theory provided a method to modify it so as to make its satisfiability more probable. Both the generation of classes of configurations and the verification of the validity of the properties being sought, however, rapidly became too complicated to be performed manually. These operations were therefore assigned to the computer, which after a very long series of attempts and calculations did identify the required class of configurations.

The spread of software for algebraic calculus, and also for integrated algebraic and numeric calculus, is an additional aspect of the development of the calculus function in theoretical physics. In all of these activities, the computer is a tool for the development

and the verification of theory. It is beyond the scope of my reflections here to discuss the advantages and the problems that result from this use of the computer.

On the other hand, the development of computers has led to the appearance of a new form of scientific research, **simulation**, which is sometimes called the *third branch* of science [Pool92], since it is placed at the same level as theory and experiment. In a simulation, the evolution equations of a system are defined to such a detail that a computer can explicitly evaluate and follow the evolution of the system. From a certain point of view, it is an experiment (one does in fact talk about numerical experiments). On the other (and this is where the most subtle dangers are hidden), the simulated system is defined exclusively by means of its evolution equations, and it remains to be seen whether and to what extent the simulated system is related to the real system.

In certain situations, the border between calculation and simulation is fairly difficult to define. Let us suppose that we want to calculate the behavior of the planets in our solar system for a certain period of time. As is well known, it is not possible to find an analytical solution to the problem of three or more gravitating bodies—however, many refined techniques for the approximate solution of the equations of motion in celestial mechanics exist. The power of computers allows to solve the problem, at least for not too long time intervals, by solving the Newton equations of motion for the planets and the sun by brute force. In this case, we can say that we are "simulating" the evolution of the planets or, that we are simply calculating the solution of the Newton equations.

These ambiguities do not occur for the systems of statistical mechanics—in this case, the distinction between simulation and calculation activities is usually fairly clear. So theory (and experiment) actually relates to macroscopic systems, which are well described by the thermodynamic limit, while simulation necessarily concerns systems of limited size—usually *very* small. The most serious technical problem one has to face, therefore, is interpreting simulation data by extrapolating the results from finite systems to "infinite" systems in the thermodynamic limit.

Numerical experiments, on the other hand, also exhibit a number of advantages. Once the model has been defined, the calculation does not contain approximations and is perfectly explicit. This is an advantage compared to the analytic treatment, which usually requires intermediate hypotheses, approximations, and so on. The results can be directly compared with real experiments so as to determine the model's validity, and with analytic solutions in order to evaluate the validity of the various approximations. It is also possible to define the model for situations that are not experimentally accessible, so as to check the validity of theoretical predictions—this is typically the case of the predictions, made on the basis of the renormalization group, that the critical exponents assume classical values for spatial dimensions higher than 4. One should add that numerical experiments give one access to microscopic details (such as each particle's position and motion at each instant) that are not available in real experiments. Last but not least, one has complete control over a numerical experiment: one does not have to worry about the sample's purity, thermostatization, and so on. As with all experimental activity, it is difficult to understand numerical simulation unless one actually carries it out. I cannot undertake an abstract discussion of the methods and the problems connected with numerical simulation here, but I refer the reader to the book by Binder and Heermann [Bind19] or the more recent one by Böttcher

and Hermann [Bott21] as possible practical guides to numerical simulation in statistical mechanics—they can also be used as sources for simulation activities to be performed on one's PC at home.

I will focus on the simulation of thermodynamic systems at equilibrium. This activity can be performed in at least two different ways: **molecular dynamics** and the **Monte Carlo method**. The specific aim of molecular dynamics is the simulation of mechanical systems (such as fluids composed of interacting particles), while the Monte Carlo method is more flexible. In molecular dynamics, one sets out to simulate the evolution of a mechanical system composed of a large number of particles, solving its equations of motion numerically. One obvious advantage is that in this fashion, one gains direct information not only about the system's static properties, but also about its dynamic properties. In the Monte Carlo method, one instead solves a fictitious evolution process, chosen to replicate the equilibrium distribution predicted by statistical mechanics. This has a series of advantages: first, it is possible to simulate systems whose dynamics are not defined (such as the Ising model), and second, it is possible to choose the fictitious dynamics so as to reach equilibrium as fast as possible, in order to optimize the quality of the data one gathers. On the other hand, while molecular dynamics can constitute an "experimental" proof of the foundations of statistical mechanics—in the sense that it allows one to verify whether the predictions of statistical mechanics follow from the solution of the equations of motion—the Monte Carlo method *presupposes* its validity, as a constraint on the choice of the fictitious dynamics.

In any case, in both situations the result of the simulation is a trajectory in the space of microscopic states, starting from which one calculates averages of the observables, correlation functions such as the $g(r)$, and so on. The problems related to the interpretation of this data are similar for both techniques.

8.2 Molecular Dynamics

In this section, we will show how one can perform the numerical simulation of a classical mechanical system by means of the molecular dynamics method. This consists of the explicit solution of Newton's equations of motion for the particles that constitute the system. It has the advantage that it allows one to evaluate at the same time both equilibrium thermodynamic properties (such as the equation of state) and dynamic properties such as viscosity, thermal conductivity, and so on. Its main disadvantage lies in the complexity of the calculations, which require significant resources even for very simple and small systems.

Let us specifically consider a model of a simple fluid, composed of N identical classical particles each of mass m, which evolve while subject to mutual interactions, described by a pair potential $u(r)$, where r is the distance between the two particles of the pair. Let us denote the momentum vector of the ith particle by $\boldsymbol{p}_i$ and its position vector by $\boldsymbol{r}_i$. The Hamiltonian of this system is then given by

$$\mathcal{H}(x) = \sum_i \frac{p_i^2}{2m} + \frac{1}{2} \sum_{i \neq j} u\left(\left|\boldsymbol{r}_i - \boldsymbol{r}_j\right|\right), \tag{8.1}$$

where $x = (\boldsymbol{p}^N, \boldsymbol{r}^N)$ is the phase-space point of the system. The Newton equations of motion for the system of N particles then read

$$m\frac{\mathrm{d}^2\boldsymbol{r}_i}{\mathrm{d}t^2} = \sum_{j(\neq i)} \boldsymbol{f}_{ij}, \tag{8.2}$$

where $\boldsymbol{f}_{ij}$ is the force with which particle j acts on particle i:

$$\boldsymbol{f}_{ij} = -\frac{\partial u\left(\left|\boldsymbol{r}_i - \boldsymbol{r}_j\right|\right)}{\partial \boldsymbol{r}_i}. \tag{8.3}$$

We can solve numerically this system of ordinary differential equations by introducing a small time increment τ. Suppose that at a given time t we are given the momenta $\boldsymbol{p}^N(t) = (\boldsymbol{p}_i(t))$ and positions $\boldsymbol{r}^N(t) = (\boldsymbol{r}_i(t))$. Then we can take advantage of the Newton equations to evaluate the momenta $\boldsymbol{p}^N(t+\tau) = (\boldsymbol{p}_i(t+\tau))$ and positions $\boldsymbol{r}^N(t+\tau) = (\boldsymbol{r}_i(t+\tau))$ at the subsequent, but close, time $t+\tau$. The procedure can then be iterated. We thus obtain a discrete approximation to the phase-space trajectory:

$$\begin{aligned}(x(t)) &\simeq (x(t_0), x(t_1), \ldots, x(t_k), \ldots) \\ &= \left(\left(\boldsymbol{p}^N(t_0), \boldsymbol{r}^N(t_0)\right), \left(\boldsymbol{p}^N(t_1), \boldsymbol{r}^N(t_1)\right), \ldots, \left(\boldsymbol{p}^N(t_k), \boldsymbol{r}^N(t_k)\right)\right), \ldots\right),\end{aligned} \tag{8.4}$$

where

$$t_k = t_0 + k\tau, \qquad k \in \{0, 1, 2, \ldots\}. \tag{8.5}$$

Given the initial conditions $\left(\boldsymbol{p}^N(t_0), \boldsymbol{r}^N(t_0)\right)$ and a sequence of $\mathcal{N}$ temporal intervals, we univocally define a sequence of $\mathcal{N}$ states in phase space. Now, if the hypotheses we made about statistical mechanics are valid, then if the number of particles N is large enough and we follow the system evolution for a duration $\mathcal{N}\tau$ of the order of the observation time that interests us, the average of the observables along this evolution must tend toward the ensemble averages. We can informally set

$$\left\langle A(x)\right\rangle = \lim_{\mathcal{N}\to\infty} \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} A\left(x(t_k)\right). \tag{8.6}$$

In the case we discussed, since the dynamics expressed by the Newton equations conserve energy, this average corresponds to the average over the microcanonical ensemble.

This conclusion should be taken with a great deal of caution. One of the first applications of computer simulation was performed by E. Fermi, J. Pasta, S.Ulam, and M. Tsingou on a one-dimensional model of coupled anharmonic oscillators [Ferm55]. To their great surprise, they saw that the system tended to return, after a long but not terribly complicated cycle, to a condition close to the one it had started from. Practically speaking, the system never "forgot" its initial state, and therefore the averages along the trajectory were limited to only one particular region of phase space. This phenomenon can appear in a variety of cases and can be difficult to identify. It usually becomes increasingly less probable as the system's size increases. It is, however, one of the simulator's tasks to verify that this phenomenon does not occur in the case being considered.

The quantities that can be obtained from a molecular dynamics trajectory are, as we saw, averages of thermodynamic observables—in other words, of functions $A(x)$ of the microstates $x = (\boldsymbol{p}^N, \boldsymbol{r}^N)$. Many interesting thermodynamic quantities are not observables in this sense: in particular, entropy, which is a property of the *macrostate* (being the logarithm of the accessible volume), and, consequently, temperature, pressure, chemical potentials, and so on, which are expressed in terms of its derivatives. In order for simulations in molecular dynamics to be of some utility, it is necessary to find a method to get around this difficulty. This can be accomplished, in the case of a classical mechanical system, by exploiting the equipartition theorem to calculate temperature and the virial theorem to calculate pressure.

8.2.1 Temperature and Pressure in Molecular Dynamics

For a classical mechanical system whose Hamiltonian is given by the sum of the kinetic energy $K(\boldsymbol{p}^N)$ and the potential energy $U(\boldsymbol{r}^N)$, where

$$K = \sum_{i=1}^{N} \sum_{\alpha=1}^{d} \frac{p_{i\alpha}^2}{2m_i}, \tag{8.7}$$

the equipartition theorem stipulates that $\langle K \rangle$ be equal to $k_B T/2$ for each degree of freedom:

$$\langle K \rangle = \frac{1}{2} N d\, k_B T. \tag{8.8}$$

This observation suggests that we calculate T by evaluating $\langle K \rangle$, which is a mechanical observable. Let us note that in this manner, we proceed to calculate T as a function of internal energy E, which, in its turn, is determined by the initial conditions. In practice, there will be difficulties when approaching the lower temperatures, because in order to do so, it is necessary to choose the initial conditions in such a way that the total energy is as small as possible. If one wants to simulate the system at low temperatures, one resorts to the trick of artificially "cooling" the system, by simultaneously dividing the velocities of all the particles by a factor $\lambda > 1$ from time to time. By applying this transformation several times, one gradually leads the system to states of smaller and smaller internal energy.

Knowledge of T as a function of the internal energy E allows one in principle to calculate the entropy S by means of integration. The specific heat C can be obtained by taking the derivative of the temperature T with respect to the internal energy E:

$$\frac{1}{C} = \frac{\partial T}{\partial E}\bigg)_N. \tag{8.9}$$

In this manner, it is also possible to obtain the entropy by integrating C/T numerically:

$$S(B) - S(A) = \int_A^B \frac{\mathrm{d}E}{T} = \int_A^B \mathrm{d}E\, \frac{1}{T(E)} \left(\frac{\partial T}{\partial E}\right)^{-1}, \tag{8.10}$$

where A and B are states of thermodynamic equilibrium. In order to be able to exploit this formula, however, it is essential to verify that the system has reached thermodynamic

equilibrium in each of the intermediate states. Otherwise, the integral will come to depend on the path.

In order to calculate the pressure, one could evaluate the mean of the force exerted by the walls on the system, consistently with the mechanical definition of pressure. This requires that we explicitly introduce the effects of the walls into the simulation. On the other hand, one usually uses periodic boundary conditions to minimize surface effects. As a consequence, walls do not appear in the actually simulated system. For this reason, the pressure is usually calculated by resorting to the **virial theorem**. The quantity

$$\Pi = \sum_i \boldsymbol{p}_i \cdot \boldsymbol{r}_i, \tag{8.11}$$

is an observable called the **virial**. At equilibrium, its average (like that of all observables) is constant. Therefore,

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} \langle \Pi \rangle = \left\langle \sum_i \frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} \cdot \boldsymbol{r}_i \right\rangle + \left\langle \sum_i \boldsymbol{p}_i \cdot \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} \right\rangle. \tag{8.12}$$

One obviously has

$$\left\langle \sum_i \boldsymbol{p}_i \cdot \frac{\mathrm{d}\boldsymbol{r}_i}{\mathrm{d}t} \right\rangle = \left\langle \sum_i \frac{p_i^2}{m_i} \right\rangle = 2 \langle K \rangle = N d \, k_\mathrm{B} T, \tag{8.13}$$

because of the equipartition theorem.

On the other hand, due to Newton's law,

$$\frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} = \boldsymbol{F}_i, \tag{8.14}$$

where $\boldsymbol{F}_i$ is the total force applied to particle i. This force is the sum of the total external forces $\boldsymbol{f}^{\mathrm{ext}}$ and of the total internal (interaction) ones $\boldsymbol{f}_i^{\mathrm{int}}$. We thus have

$$\left\langle \sum_i \frac{\mathrm{d}\boldsymbol{p}_i}{\mathrm{d}t} \cdot \boldsymbol{r}_i \right\rangle = \sum_i \left\langle \boldsymbol{f}_i^{\mathrm{ext}} \cdot \boldsymbol{r}_i \right\rangle + \sum_i \left\langle \boldsymbol{f}_i^{\mathrm{int}} \cdot \boldsymbol{r}_i \right\rangle. \tag{8.15}$$

If the external force is applied by the walls, and the pressure is equal to p, one has

$$\sum_i \left\langle \boldsymbol{f}_i^{\mathrm{ext}} \cdot \boldsymbol{r}_i \right\rangle = -p \oint \mathrm{d}A \, \boldsymbol{n} \cdot \boldsymbol{r}, \tag{8.16}$$

where $\mathrm{d}A$ is the area element of the walls, $\boldsymbol{n}$ is the corresponding normal unit vector (oriented toward the outside), and the integral is extended to the sample's external surface. By Gauss's theorem,

$$\oint \mathrm{d}A \, \boldsymbol{n} \cdot \boldsymbol{r} = \int \mathrm{d}\boldsymbol{r} \; (\nabla \cdot \boldsymbol{r}) = d \, V, \tag{8.17}$$

where V is the sample volume and d the space dimensionality. We thus obtain

$$\sum_i \left\langle \boldsymbol{f}_i^{\text{int}} \cdot \boldsymbol{r}_i \right\rangle = -d\, p\, V. \tag{8.18}$$

This yields the following expression of the pressure:

$$pV = N k_{\text{B}} T + \frac{1}{d} \left\langle \sum_i \boldsymbol{f}_i^{\text{int}} \cdot \boldsymbol{r}_i \right\rangle. \tag{8.19}$$

Since the calculation of the dynamics requires the evaluation of the $\boldsymbol{r}_i$ and $\boldsymbol{f}_i^{\text{int}}$ at each time step, it is possible to evaluate the pressure directly by straightforward averaging.

This expression is simplified if the internal forces are due only to pair interactions—in other words, if

$$\boldsymbol{f}_i^{\text{int}} = \sum_{j(\neq i)} \boldsymbol{f}_{ij}, \tag{8.20}$$

where $\boldsymbol{f}_{ij}$ is the force that particle j applies on particle i and satisfies $\boldsymbol{f}_{ij} = -\boldsymbol{f}_{ji}$. We can then group the contribution of each pair of particles into the right-hand side of the eq. (8.19), obtaining

$$\sum_i \sum_{j(\neq i)} \left\langle \boldsymbol{f}_{ij} \cdot \boldsymbol{r}_i \right\rangle = \frac{1}{2} \sum_{i \neq j} \left\langle \boldsymbol{f}_{ij} \cdot \boldsymbol{r}_{ij} \right\rangle, \tag{8.21}$$

where $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$. By grouping all the terms, we thus obtain

$$pV = N k_{\text{B}} T + \frac{1}{2d} \sum_{i \neq j} \langle \boldsymbol{f}_{ij} \cdot \boldsymbol{r}_{ij} \rangle = N k_{\text{B}} T - \frac{1}{2d} \sum_{i \neq j} \left\langle \boldsymbol{r}_{ij} \cdot \frac{\partial u(r_{ij})}{\partial \boldsymbol{r}_{ij}} \right\rangle. \tag{8.22}$$

Exercise 8.1 (Ideal gas in one dimension) Consider a system composed of N particles of mass m in one dimension, contained in the half-line $r_i \geq 0$, $i = 1, 2, \ldots, N$, with an elastic wall at the origin. The particles are confined by a piston of mass M, whose instantaneous position is given by $L(t) > 0$, and which is acted on by a force p oriented toward the origin.

Solve this problem in molecular dynamics, taking into account the fact that the particles interact only with the fixed wall at the origin (where they bounce elastically) and with the piston, by means of impacts in which energy and momentum are conserved. Check the conservation of enthalpy $E + pL$. Show (numerically) that the velocity distribution tends to the Maxwell distribution and that the equation of state is given approximately by the law of ideal gases.

8.2.2 Verlet Algorithms

One of the most frequently used algorithms for solving numerically Newton's equations is due to L. Verlet [Verl67]. Let us take the case of one degree of freedom that satisfies the

equation of motion:

$$\ddot{x}(t) = f(x(t)), \tag{8.23}$$

where the dot denotes the derivative with respect to time. By expanding $x(t+\tau)$ in a series around t, we obtain

$$x(t+\tau) = x(t) + \tau\,\dot{x}(t) + \frac{1}{2}\tau^2\,\ddot{x}(t) + \frac{1}{6}\tau^3\,\dddot{x}(t) + \mathrm{O}(\tau^4). \tag{8.24}$$

By summing with the corresponding expression for $x(t-\tau)$, we obtain

$$x(t+\tau) + x(t-\tau) = 2x(t) + \tau^2\,\ddot{x}(t) + \mathrm{O}(\tau^4), \tag{8.25}$$

and therefore, taking into account eq. (8.23),

$$x(t+\tau) = 2x(t) - x(t-\tau) + \tau^2 f(x(t)) + \mathrm{O}(\tau^4). \tag{8.26}$$

This is a third-order algorithm (in other words, it disregards terms on the order of τ^4), and has the advantage of being symmetrical with respect to time reversal. It is, moreover, very simple to put into practice. One needs, however, to keep in mind that the method requires us to know $x(t)$ and the preceding value $x(t-\tau)$ at each instant. If, as usually occurs, the initial positions and velocities are known, it is necessary to estimate the $x(t-\tau)$ by solving one step of the differential equation with some other method. This often introduces an error on the order of τ^3. This does not create difficulties in the case of statistical averages, but it can do so if one is looking for the actual solution of the differential equation for a given initial condition. Moreover, in order to evaluate the energy, it is necessary to estimate the velocity at each time step. It is easy to see that a good estimator of the velocity at time t is obtained by the symmetric difference

$$v(t) = \frac{x(t+\tau) - x(t-\tau)}{2\tau} = \dot{x}(t) + \mathrm{O}(\tau^2). \tag{8.27}$$

Exercise 8.2 Write the Verlet algorithm for the pendulum

$$\ell\,\ddot{\theta} = -g\sin\theta.$$

1. Solve the equations of motion for $\theta(t)$ numerically by the Verlet algorithm, and check energy conservation.
2. Experiment with the initial conditions, and evaluate numerically how the oscillation period grows when the initial angle approaches 180° (supposing that it starts at rest).

One can obtain a more precise evaluation of the velocity at each time step by the **velocity Verlet** method. This method formally involves two evaluations of $f(x(t))$ per step, but each evaluation is used in two successive steps, so that the required computational effort is essentially the same as in the standard Verlet method. Suppose we know $x(t)$ and $v(t)$. Let

us estimate $v(t+\tau/2)$:

$$v(t+\tau/2) = v(t) + (\tau/2) f(x(t)) + \mathrm{O}(\tau^2). \tag{8.28}$$

To evaluate $x(t+\tau)$, we use the approximation

$$x(t+\tau) \simeq x(t) + \tau v(t+\tau/2). \tag{8.29}$$

To evaluate $v(t+\tau)$, we use again (8.28):

$$v(t+\tau) \simeq v(t+\tau/2) + (\tau/2) f(x(t+\tau)). \tag{8.30}$$

Eliminating $v(t+\tau/2)$ by eq. (8.28), we obtain the following algorithm:

$$x(t+\tau) \simeq x(t) + \tau v(t) + (\tau^2/2) f(x(t)); \tag{8.31a}$$

$$v(t+\tau) \simeq v(t) + (\tau/2) \left(f(x(t)) + f(x(t+\tau)) \right). \tag{8.31b}$$

This method yields an update of the *pair* (x, v) at each step. Note that the second evaluation of the force that appears in the velocity equation can be used for the next update of $x(t)$. One can show that the errors on $v(t+\tau)$ and $x(t+\tau)$ are of order τ^4, such as the direct Verlet algorithm.

In the case of simple fluid models, it is necessary to choose the initial condition with some care. One usually prefers to initially arrange the particles on a regular lattice, so as to guarantee a given value of the density, and to assign a velocity, chosen randomly from a Gaussian distribution, to each of them. By changing the variance of the Gaussian, one can change the value of the temperature.

Exercise 8.3 Consider a one-dimensional fluid defined by the Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i=1}^{N} u(r_{i+1} - r_i),$$

with periodic boundary conditions:

$$r_{i+N} = r_i, \qquad -L/2 \le r_i \le L/2, \qquad u(r_i + L) = u(r_i), \qquad \forall i.$$

Let $u(r)$ have the Lennard-Jones expression:

$$u(r) = \epsilon_0 \left[\left(\frac{r_0}{r} \right)^{12} - \left(\frac{r_0}{r} \right)^{6} \right],$$

where it is assumed that only nearest-neighbor particles interact.

1. Introduce suitable length and time scales to make the differential equation adimensional. This leaves an explicit dependence on the adimensional particle density.

2. Write the program for the solution of the equations of motion by means of the velocity Verlet algorithm.

3. Write the expressions of the temperature and the pressure as functions of the calculated trajectories obtained by means of the Verlet algorithm.

4. Evaluate the order of magnitude (in real length and time) of the systems and durations you can simulate on your PC, assuming $m \sim 10^{-23}$ g, $\epsilon_0 \sim 1$ eV, and $r_0 \sim 1$ Å.

5. Choose the initial conditions so that the r_i are arranged on a regular lattice and the p_i follow a Gaussian distribution. Launch the simulation. Check the conservation of energy and monitor the approach to equilibrium. You can also compare your data to the exact solution. Be careful not to use too small values for the energy, because the Fermi-Pasta-Ulam-Tsingou phenomenon mentioned earlier can make the approach to equilibrium in finite systems unreliable and slow.

The success of the method of molecular dynamics in the study of the statistical mechanics of fluids represents one of the most powerful arguments for the validity of the entire discipline. In fact, it guarantees that the main results of statistical mechanics descend directly from the properties of the equations of motion in the case of a fairly large system of particles, and it is reassuring to know that the required size of the system is not excessive. (Several thousand particles can be sufficient for simple systems.) One often has to deal, however, with systems for which the molecular dynamics approach is not warranted: either the evolution of the system is not described by ordinary differential equations, or solving these equations becomes too burdensome. In this case, it is convenient to directly sample the microstate distribution defined by the statistical ensemble. This can be obtained by the Monte Carlo method, which we discuss later in this chapter.

8.3 Thermostats in Molecular Dynamics

When simulating a particle system via Molecular Dynamics, we end up sampling the microcanonical ensemble. Since the system has necessarily a finite (and quite small) size, ensemble equivalence is not warranted. We can thus ask ourselves if it is possible to use molecular dynamics to sample the canonical ensemble (or even the pT ensemble), keeping its advantages, those of its being a *deterministic* and *reversible* dynamics.

In principle, one can impose a temperature T on a system S by letting it interact with a much larger reservoir R. Clearly, this is not a clever strategy, because it requires simulating a resulting system with $N = N^{(R)} + N^{(S)}$ degrees of freedom, with $N \gg N^{(S)}$, when we are only interested in a system with $N^{(S)}$ degrees of freedom.

In one of the first thermostatting algorithms proposed for molecular dynamics, one chooses at random intervals (Poisson-distributed with a frequency ν, that is, with a probability $\nu \mathrm{d}t$ of occurring in an interval of duration $\mathrm{d}t$) one particle at random, and assigns to it a velocity extracted from a Boltzmann distribution at the given temperature [Ande80]. In principle, ν can be chosen arbitrarily. In practice, there is an optimum choice, namely

$$\nu = \frac{2a\kappa}{3k_B \rho^{1/3} N^{1/3}}, \tag{8.32}$$

where κ is the thermal conductivity, ρ the number density, N the number of particles, and a is a dimensionless constant. This method obviously violates energy conservation and the Liouville theorem.

Quite surprisingly, it is possible to sample the canonical distribution in molecular dynamics by adding to the system's degrees of freedom *a single* mechanical degree of freedom (or, more generally, N' degrees of freedom, with $N' \ll N$). This comes at the price of a rather unintuitive representation of the dynamics of the extended system. The method was introduced by S. Nosé in 1984 and perfected by W. G. Hoover the following year, and is therefore known as the **Nosé-Hoover thermostat**.

Let us thus consider a mechanical system of N particles of mass m in d dimensions, described by the Hamiltonian

$$\mathcal{H}_0\left(\boldsymbol{p}^N, \boldsymbol{r}^N\right) = \sum_{i=1}^{N} \frac{p_i^2}{2m} + U\left(\boldsymbol{r}^N\right). \tag{8.33}$$

The temperature T of such a system can be obtained by the relation

$$\left\langle \sum_{i=1}^{N} \frac{p_i^2}{2m} \right\rangle = \frac{N_{\mathrm{df}}}{2} k_{\mathrm{B}} T, \tag{8.34}$$

where $N_{\mathrm{df}} = dN - N_{\mathrm{c}}$ is the number of degrees of freedom of the system, taking into account the N_{c} constraints applied to it. For instance, in a molecular dynamics simulation with periodic boundary conditions, the total energy and the total momentum, but not the total angular momentum, are conserved. We then have $N_{\mathrm{c}} = d + 1$. Then, once one has a solution $R(t) = ((\boldsymbol{r}_i(t)))$ of the evolution equations, one can use eq. (8.34) to express T by substituting the relation

$$\boldsymbol{p}_i = m\dot{\boldsymbol{r}}_i, \qquad i = 1, 2, \ldots, N. \tag{8.35}$$

Thus one could keep the temperature constant, in principle, by rescaling the velocities $\dot{\boldsymbol{r}}_i$ by a suitably chosen common factor.

The idea of the Nosé-Hoover thermostat [Nose84, Hoov85] is to introduce an extra *mechanical* variable s' endowed with a "mass" $Q > 0$. The variable s' acts as a time rescaling factor (thus rescaling the velocities). The increment of the time coordinate t' of the extended system S′ (which contains the extra degree of freedom) is rescaled by a factor s' with respect to the increment of the physical system:

$$\mathrm{d}t' = s'\,\mathrm{d}t. \tag{8.36}$$

On the other hand, the positions of the particles do not change in going from the extended to the physical system. This is expressed by the relations

$$\boldsymbol{r}_i' = \boldsymbol{r}_i, \qquad \dot{\boldsymbol{r}}_i' = \frac{\dot{\boldsymbol{r}}_i}{s'}, \qquad i = 1, 2, \ldots, N; \tag{8.37}$$

$$s' = s, \qquad \dot{s}' = \frac{\dot{s}}{s}. \tag{8.38}$$

The Hamiltonian of the extended system is given by

$$\mathcal{H}(\{\boldsymbol{p}_i'\},\{\boldsymbol{r}_i'\},s',P') = \mathcal{H}_0 + \frac{P'^2}{2Q} + gk_{\mathrm{B}}T_0 \ln s', \tag{8.39}$$

where $\boldsymbol{p}_i' = m\boldsymbol{r}_i'$, $i = 1,\ldots,N$, one sets $g = N_{\mathrm{df}} + 1$, and the "mass" $Q > 0$ must be chosen with care. We thus obtain the evolution equations

$$\ddot{\boldsymbol{r}}_i' = \frac{\boldsymbol{f}_i'}{m\,s'^2} - \frac{2\dot{s}'\,\dot{\boldsymbol{r}}_i'}{s'}, \qquad i = 1,2,\ldots,N; \tag{8.40}$$

$$\ddot{s}' = \frac{1}{Q\,s'}\left(\sum_{i=1}^{N} m s'^2 (\dot{\boldsymbol{r}}_i')^2 - gk_{\mathrm{B}}T_0\right). \tag{8.41}$$

These equations can be solved by the velocity Verlet method. They correspond to a microcanonical ensemble for the extended system. The energy of the physical system is not conserved and the degrees of freedom associated with s' play the role of a heat reservoir. This property is proved in sec. 8.3.1. These equations are deterministic and reversible. Since the equation that governs the behavior of s' is of second order, it may exhibit oscillations, where the energy flows periodically in and out of the physical system, corresponding to periodic oscillations of the instantaneous temperature.

Since in the Nosé thermostat the time flows at a fluctuating pace, it is quite hard to interpret the simulation results. However, Hoover [Hoov85], has shown that it is possible to recast the Nosé evolution equations in terms of the variables of the physical system. This is made by the transformations

$$s = s'; \qquad \dot{s} = s'\dot{s}'; \qquad \ddot{s} = s'^2\ddot{s}' + s'(\dot{s}')^2; \tag{8.42}$$

$$\boldsymbol{r}_i = \boldsymbol{r}_i'; \qquad \dot{\boldsymbol{r}}_i = s'\dot{\boldsymbol{r}}_i; \qquad \ddot{\boldsymbol{r}}_i = s'^2\ddot{\boldsymbol{r}}_i' + s'\dot{s}'\dot{\boldsymbol{r}}_i; \qquad i = 1,\ldots,N. \tag{8.43}$$

Defining then γ by

$$\gamma = \frac{\dot{s}}{s}, \tag{8.44}$$

the equations of motion take the form

$$\ddot{\boldsymbol{r}}_i = \frac{\boldsymbol{f}_i}{m} - \gamma\dot{\boldsymbol{r}}_i; \tag{8.45}$$

$$\dot{\gamma} = \frac{1}{Q}\left(\sum_{i=1}^{N} m\dot{\boldsymbol{r}}_i^2 - g\,k_{\mathrm{B}}T_0\right). \tag{8.46}$$

The coefficient γ can take on negative values, and changes as the system evolves. We show in sec. 8.3.1 that it is necessary to set $g = N_{\mathrm{df}} + 1$ to obtain the correct dynamical sampling at the temperature T_0.

The Nosé-Hoover thermostat can be implemented by a reversible algorithm due to Martyna et al. [Mart95]. Consider the evolution of the generic variable $A(x,\gamma)$:

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \sum_{i=1}^{N}\left(\dot{\boldsymbol{r}}_i\cdot\frac{\partial A}{\partial \boldsymbol{r}_i} + \dot{\boldsymbol{p}}_i\cdot\frac{\partial A}{\partial \boldsymbol{p}_i}\right) + \dot{\gamma}\frac{\partial A}{\partial\gamma} = \mathcal{L}\,A, \tag{8.47}$$

where $\mathcal{L}$ is the Liouville operator. Formally solving this equation we obtain

$$A(x(t+\tau)) = \mathrm{e}^{\tau\mathcal{L}} A(x(t)). \tag{8.48}$$

Let us split the Liouville operator into three terms:

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{N}\left[\frac{\boldsymbol{p}_i}{m}\cdot\frac{\partial}{\partial \boldsymbol{r}_i} + \left(\boldsymbol{f}_i - \gamma \boldsymbol{p}_i\right)\cdot\frac{\partial}{\partial \boldsymbol{p}_i}\right] + \frac{1}{Q}\left(\sum_{i=1}^{N}\frac{\dot{r}_i^2}{m} - g\,k_{\mathrm{B}}T_0\right)\frac{\partial}{\partial\gamma} \\ &= \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3;\end{aligned} \tag{8.49}$$

$$\mathcal{L}_1 = \sum_{i=1}^{N}\frac{\boldsymbol{p}_i}{m}\cdot\frac{\partial}{\partial \boldsymbol{r}_i} + \frac{1}{Q}\left(\sum_{i=1}^{N}\frac{\dot{r}_i^2}{m} - g\,k_{\mathrm{B}}T_0\right)\frac{\partial}{\partial\gamma}; \tag{8.50}$$

$$\mathcal{L}_2 = \sum_{i=1}^{N}\boldsymbol{f}_i\cdot\frac{\partial}{\partial \boldsymbol{p}_i}; \tag{8.51}$$

$$\mathcal{L}_3 = -\sum_{i=1}^{N}\gamma\boldsymbol{p}_i\cdot\frac{\partial}{\partial \boldsymbol{p}_i}. \tag{8.52}$$

On the other hand, by Trotter's formula,

$$\exp\left[\tau\,(A+B)\right] = \exp\left(\frac{\tau B}{2}\right)\exp\left(\tau A\right)\exp\left(\frac{\tau B}{2}\right) + \mathrm{o}\left(\tau^2\right), \tag{8.53}$$

which can be directly verified by Taylor expansion, we have

$$\mathrm{e}^{\tau\mathcal{L}} = \mathrm{e}^{\tau\mathcal{L}_3/2}\mathrm{e}^{\tau\mathcal{L}_2/2}\mathrm{e}^{\tau\mathcal{L}_1}\mathrm{e}^{\tau\mathcal{L}_2/2}\mathrm{e}^{\tau\mathcal{L}_3/2} + \mathrm{o}\left(\tau^2\right). \tag{8.54}$$

Now

$$\mathrm{e}^{\tau\mathcal{L}_1}\boldsymbol{r}_i = \boldsymbol{r}_i + \frac{\boldsymbol{p}_i}{m}\tau; \tag{8.55}$$

$$\mathrm{e}^{\tau\mathcal{L}_1}\gamma = \gamma + \frac{\tau}{Q}\left(\sum_{i=1}^{N}\frac{p_i^2}{m} - gk_{\mathrm{B}}T_0\right). \tag{8.56}$$

We have, moreover,

$$\mathrm{e}^{\tau\mathcal{L}_2}\boldsymbol{p}_i = \boldsymbol{p}_i + \tau\boldsymbol{f}_i; \tag{8.57}$$

$$\begin{aligned}\mathrm{e}^{\tau\mathcal{L}_3}\boldsymbol{p}_i &= \left(1 + \tau\mathcal{L}_3 + \frac{\tau^2}{2}\mathcal{L}_3^2 + \cdots\right)\boldsymbol{p}_i \\ &= \boldsymbol{p}_i - \tau\,\gamma\boldsymbol{p}_i + \frac{(-\tau\gamma)^2}{2}\boldsymbol{p}_i + \cdots \\ &= \mathrm{e}^{-\tau\gamma}\boldsymbol{p}_i.\end{aligned} \tag{8.58}$$

We thus obtain the following reversible algorithm:

1. $\boldsymbol{p}_i \longleftarrow \mathrm{e}^{-\tau\gamma/2}\boldsymbol{p}_i$;
2. $\boldsymbol{p}_i \longleftarrow \boldsymbol{p}_i + \tau\boldsymbol{f}_i/2$;

3. $\boldsymbol{r}_i \longleftarrow \boldsymbol{r}_i + \tau \boldsymbol{p}_i/m$;
4. $\gamma \longleftarrow \gamma + \tau \left[\sum_{i=1}^{N} (p_i^2/m) - g\, k_B T_0\right]$;
5. $\boldsymbol{p}_i \longleftarrow \boldsymbol{p}_i + \tau \boldsymbol{f}_i/2$;
6. $\boldsymbol{p}_i \longleftarrow \mathrm{e}^{-\tau\gamma/2}\boldsymbol{p}_i$.

Using similar ideas, Martyna and coworkers were able to design an algorithm for the simulation of the pT ensemble in molecular dynamics [Mart94].

In the study of complex molecular systems, a different kind of approach is often used, known as **dissipative particle dynamics** (DPD) [Hoog92] (see the review in [Sant21]). Here one adds to the conservative forces that give the system its chemical identity a set of dissipative (deterministic) and random forces chosen to satisfy local momentum conservation and to keep the mean temperature of the system constant. Since discussing this technique requires some results that will be obtained in the next chapter, I will not dwell on it further.

8.3.1 Proof of the Canonical Sampling

The simulation of the extended system S′ in molecular dynamics samples the correspondent microcanonical ensemble. Consider the expression of the corresponding partition function:

$$Z = \int \mathrm{d}s'\, \mathrm{d}P' \int \prod_{i=1}^{N} (\mathrm{d}\boldsymbol{p}_i'\, \mathrm{d}\boldsymbol{r}_i')\, \delta\left(\mathcal{H}(\{\boldsymbol{p}_i'\}, \{\boldsymbol{r}_i'\}, s', P') - E\right). \tag{8.59}$$

Let us go from virtual to physical variables by means of eq. (8.37). We then get the following expression of Z:

$$Z = \int \mathrm{d}s'\, \mathrm{d}P'\, s'^{N_{\mathrm{df}}} \int \mathrm{d}\boldsymbol{p}^N\, \mathrm{d}\boldsymbol{r}^N\, \delta\left(\mathcal{H}_0\left(\boldsymbol{p}^N, \boldsymbol{r}^N\right) + \frac{P^2}{2Q} + g\, k_B T_0 \ln s' - E\right). \tag{8.60}$$

By exploiting the following identity of the delta function (see sec. A.8)

$$\delta\left(f(s')\right) = \frac{1}{\left|f'(s')\right|}\delta(s' - s_0), \tag{8.61}$$

where $f(s_0) = 0$, we obtain

$$s_0 = \exp\left[\frac{1}{g\, k_B T_0}\left(E - \mathcal{H}_0\left(\boldsymbol{p}^N, \boldsymbol{r}^N\right) - \frac{P^2}{2Q}\right)\right]; \tag{8.62}$$

$$\begin{aligned} Z &= \int \mathrm{d}s'\, \mathrm{d}P'\, s'^{N_{\mathrm{df}}} \int \mathrm{d}\boldsymbol{p}^N\, \mathrm{d}\boldsymbol{r}^N\, \frac{\delta(s' - s_0)}{g k_B T_0/s'} \\ &= \int \mathrm{d}P' \int \mathrm{d}\boldsymbol{p}^N\, \mathrm{d}\boldsymbol{r}^N\, \frac{1}{g\, k_B T_0} \exp\left[\frac{N_{\mathrm{df}} + 1}{k_B T_0}\left(E - \mathcal{H}_0 - \frac{P'^2}{2Q}\right)\right] \end{aligned}$$

$$= \int \mathrm{d}\boldsymbol{p}^N \, \mathrm{d}\boldsymbol{r}^N \, \exp\left[-\frac{N_{\mathrm{df}}+1}{g\, k_{\mathrm{B}} T_0} \mathcal{H}_0 \left(\boldsymbol{p}^N, \boldsymbol{r}^N \right) \right]$$

$$\times \int \mathrm{d}P' \, \frac{1}{g\, k_{\mathrm{B}} T_0} \exp\left[\frac{1}{k_{\mathrm{B}} T_0} \left(E - \frac{P'^2}{2Q} \right) \right]. \tag{8.63}$$

The factor in the last line does not depend on $\left(\boldsymbol{p}^N, \boldsymbol{r}^N\right)$, and changes therefore the normalization but not the distribution. Letting $g = N_{\mathrm{df}} + 1$ we obtain the canonical distribution at the temperature T_0.

8.4 Monte Carlo Method

Let us suppose that we want to calculate an integral such as

$$\langle A \rangle = \int_0^1 \mathrm{d}x \, A(x) \, \rho(x), \tag{8.64}$$

where $A(x)$ is a function of x, and $\rho(x) \geq 0$ is such that $\int_0^1 \mathrm{d}x \, \rho(x) = 1$, and can thus be interpreted as a probability distribution. The most straightforward way of numerically evaluating this integral is to evaluate $A(x)\rho(x)$ at $\mathcal{N}+1$ points uniformly placed between 0 and 1, $x_k = k/\mathcal{N}$, $k = 0, 1, \ldots, \mathcal{N}$, and to set

$$\langle A \rangle \simeq \overline{A} = \frac{1}{\mathcal{N}+1} \sum_{k=0}^{\mathcal{N}} A(x_k) \rho(x_k). \tag{8.65}$$

If necessary, we can improve on this estimate by using Simpson's rule or similar methods.

This method is less effective, however, if the integrand, and $\rho(x)$ in particular, is much larger in some regions of the interval than in others. In fact, the regions in which $\rho(x) \simeq 0$ give a negligible contribution to $\langle A \rangle$. If $A(x)$ is a "smooth" function, one gets a better estimate of $\langle A \rangle$ by distributing the x_k so that $\rho(x_k)\Delta x_k = \rho(x_k)\,(x_{k+1} - x_k)$ is approximately constant. It is easy to obtain such a distribution in one dimension by introducing the cumulative distribution $\Phi(x)$:

$$\Phi(x) = \int_0^x \mathrm{d}x' \, \rho(x'). \tag{8.66}$$

By then choosing the x_k such that

$$\Phi(x_{k+1}) - \Phi(x_k) = \frac{1}{\mathcal{N}}, \tag{8.67}$$

we obtain the following approximation of $\langle A \rangle$:

$$\langle A \rangle \simeq \frac{1}{\mathcal{N}+1} \sum_{k=0}^{\mathcal{N}} A(x_k). \tag{8.68}$$

In this expression, all points give approximately the same contribution, at least if $A(x)$ is a slowly varying function.

This procedure is quite effective in one dimension, although it requires one to solve eq. (8.67), which may need some effort. It has the shortcoming, however, that it cannot be efficiently adapted to cases where x varies in a space with more than a few dimensions.

Suppose for instance that we are dealing with a d-dimensional integral:

$$\langle A \rangle = \int \mathrm{d}\boldsymbol{x}\, A(\boldsymbol{x})\rho(\boldsymbol{x}), \tag{8.69}$$

where $\boldsymbol{x} = (x^\alpha)$, $0 \leq x^\alpha \leq 1$, $\alpha = 1, \ldots, d$. In order to obtain a uniform distribution of the $\boldsymbol{x}_k$, we need a number of points that grows exponentially with d. This number can rapidly become too large even for moderate values of d. On the other hand, if we consider the expression

$$\overline{A} = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} A(\boldsymbol{x}_k)\rho(\boldsymbol{x}_k), \tag{8.70}$$

where the $\boldsymbol{x}_k$ are random, independent, and identically distributed vectors (e.g., uniformly distributed in the unit hypercube in d dimensions), then, as a result of the central limit theorem, $\overline{A}$ is a Gaussian variable, whose average is equal to $\langle A \rangle$ and whose variance is proportional to $\mathcal{N}^{-1}$. We can therefore use $\overline{A}$ as an estimate of $\langle A \rangle$, expecting a random deviation ("statistical error") of the order of $\mathcal{N}^{-1/2}$. This approach to the asymptotic value is very slow (in order to improve the estimate by one significant digit, one needs 100 times more points), but its behavior is independent of d.

In order to improve the convergence, it is obviously worth our while to avoid having many points $\boldsymbol{x}_k$ falling into regions where $\rho(\boldsymbol{x})$ is small. As we saw earlier, the best situation is that in which the density of the points is proportional to $\rho(\boldsymbol{x})$, so that the number of points contained in a small volume area, centered at $\boldsymbol{x}$ and of volume $\mathrm{d}\boldsymbol{x}$, is given by $\mathcal{N}\,\rho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. In this situation, we have

$$\langle A \rangle \simeq \overline{A} = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} A(\boldsymbol{x}_k), \tag{8.71}$$

with a statistical error on the order of $\mathcal{N}^{-1/2}$.

The idea of the Monte Carlo method is to evaluate the average (8.69) of a quantity $A(\boldsymbol{x})$ over a distribution $\rho(\boldsymbol{x})$ using expression (8.71), where the $\boldsymbol{x}_k$ are random, independent, and distributed with a probability density $\rho(\boldsymbol{x})$.

One should immediately add that this "ideal" program cannot be realized. Except for trivial cases, it is not possible to find a procedure that allows one to extract some $\boldsymbol{x}$s with a probability distribution $\boldsymbol{x}$, independently of one another. The central limit theorem, however, guarantees us that $\langle A \rangle = \overline{A} + \mathrm{O}\left(\mathcal{N}^{-1/2}\right)$ even if the $\boldsymbol{x}_k$ are *not* mutually independent. It is sufficient that the $(\boldsymbol{x}_k)$ sequence be such that:

1. the distribution of each of the $\boldsymbol{x}_k$ is equal to $\rho(\boldsymbol{x})$,

2. the correlation between $\boldsymbol{x}_k$ and $\boldsymbol{x}_{k+\ell}$ goes to zero fairly rapidly as $|\ell|$ grows.

The problem thus boils down to finding an algorithm that extracts the $\boldsymbol{x}_k$ with probability distribution $\rho(\boldsymbol{x})$, in such a way that correlations among the $\boldsymbol{x}_k$ along the sequence go to zero fairly rapidly. This problem can be solved by introducing a suitable **Markov chain**.

8.4.1 Markov Chains

To keep things simple, we consider a system whose states are discrete, and belong to a set Q. The **Markov chain** X is a sequence of random variables X_t, $t = 0, 1, 2, \ldots$, such that

1. $X_t \in Q$, $\forall t$.
2. The conditional probability distribution of X_{t+1} over Q, given the value x_t of X_t, depends only on x_t and not on the values of other $X_{t'}$, with $t' \neq t, t+1$.

In other words, in a Markov chain, the probability distribution of X_{t+1} depends only on the state of X_t and not on the preceding (or following!) ones.

The essential properties of Markov chains are summarized in Appendix G. It is clear from the points listed above that the statistical properties of the sequence (X_t) are determined by the conditional probabilities $W_{xx'}(t) = P(X_{t+1} = x | X_t = x')$, $\forall x, x' \in Q$, called **transition probabilities**, and by the initial probability distribution $p_x^0 = P(X_0 = x)$, $\forall x \in Q$. We only consider here the case of time-independent transition probabilites $\mathbb{W} = (W_{xx'})$.

Some important properties follow from the interpretation of the $W_{xx'}$ as transition probabilities:

1. $W_{xx'} \geq 0$, $\forall x, x' \in Q$.
2. $\sum_{x \in Q} W_{xx'} = 1$, $\forall x' \in Q$.

Matrices $\mathbb{W} = (W_{xx'})$ that satisfy these conditions are called **stochastic**.

Given the matrix $\mathbb{W}$ of the transition probabilities, also called the **transition matrix**, the probability $p_x(t) = P(X_t = x)$ satisfies the evolution equation

$$p_x(t+1) = \sum_{x'} W_{xx'} p_{x'}(t). \tag{8.72}$$

By exploiting property 2 of stochastic matrices, this equation can be put in the **master equation** form (G.3):

$$\Delta p_x(t) = p_x(t+1) - p_x(t) = \sum_{x' (\neq x)} \left[W_{xx'} p_{x'}(t) - W_{x'x} p_x(t) \right]. \tag{8.73}$$

We want to look for the stationary distribution $p^{\mathrm{st}} = (p_x^{\mathrm{st}})$ of the X_t—in other words, a distribution such that if $p_x(t) = p_x^{\mathrm{st}}$, $\forall x$, at instant t, it remains equal to p_x^{st} for all following instants. In Appendix G it is shown that, if the transition matrix $\mathbb{W}$ is **ergodic**, i.e., if there is a integer s such that $\mathbb{W}^n$ has all positive entries for $n \geq s$, such a distribution is unique and is reached exponentially fast as $t \to \infty$. Clearly, such a distribution must satisfy the equation

$$\sum_{x' (\neq x)} \left[W_{xx'} p_{x'}^{\mathrm{st}} - W_{x'x} p_x^{\mathrm{st}} \right] = 0, \qquad \forall x. \tag{8.74}$$

A sufficient condition for this equation to hold is that the terms in brackets cancel out for each pair (x, x') of different states:

$$W_{xx'} p^{\mathrm{st}}_{x'} - W_{x'x} p^{\mathrm{st}}_{x} = 0, \qquad \forall x \neq x'. \tag{8.75}$$

In this case the outflow from state x to state x' is exactly cancelled by the inflow from state x' to state x. In this situation, one says that the stationary state satisfies **detailed balance**, and one often denotes the stationary distribution as the **equilibrium distribution** $p^{\mathrm{eq}} = (p^{\mathrm{eq}}_x)$.

If the stationary distribution has the detailed balance property, the transition matrix must be such that, for any sequence $(x_0, x_1, \dots, x_r)$ of distinct states, one has

$$\frac{W_{x_0x_1} W_{x_1x_2} \cdots W_{x_rx_0}}{W_{x_0x_r} W_{x_rx_{r-1}} \cdots W_{x_1x_0}} = \frac{W_{x_0x_1}}{W_{x_1x_0}} \frac{W_{x_1x_2}}{W_{x_2x_1}} \cdots \frac{W_{x_rx_0}}{W_{x_0x_r}} = \frac{p^{\mathrm{eq}}_{x_0}}{p^{\mathrm{eq}}_{x_1}} \frac{p^{\mathrm{eq}}_{x_1}}{p^{\mathrm{eq}}_{x_2}} \cdots \frac{p^{\mathrm{eq}}_{x_r}}{p^{\mathrm{eq}}_{x_0}} = 1. \tag{8.76}$$

In other words, for any such sequence, the product of the matrix elements obtained by looping the sequence in one direction must be equal to the product obtained by looping in the reverse direction. This requires in particular that, if $W_{xx'} > 0$ for some x and x', one must also have $W_{x'x} > 0$.

If the matrix W satisfies eq. (8.76) for all sequences of distinct states, the stationary distribution satisfies detailed balance and it is possible to obtain it explicitly. Choose any reference state x_0 and assign to it a weight $w_{x_0} > 0$. Assign to any state x_1 such that $W_{x_1x_0} > 0$ the weight $w_{x_1} = W_{x_1x_0} w_{x_0} / W_{x_0x_1}$. Now look for all states x_2 such that at least one $W_{x_2x_1} > 0$, and assign each of them the weight $w_{x_2} = W_{x_2x_1} w_{x_1} / W_{x_1x_2}$. Iterate the procedure until all states x are covered. The detailed balance property will make it so that if a state x is reached from x_0 via two different paths, it will receive the same weight.

Exercise 8.4 Prove the preceding assertion.

Once the weights have been assigned to every state in Q, the equilibrium probability is given by

$$p^{\mathrm{eq}}_x = \frac{w_x}{\sum_{x'} w_{x'}}, \qquad \forall x. \tag{8.77}$$

It must be emphasized, however, that detailed balance is by no means a necessary condition for stationarity. If we are able to find an ergodic matrix W of transition probabilities that satisfies the stationarity condition (8.74) with the desired stationary distribution p^{st}, the results of Appendix G guarantee that a Markov chain with that transition matrix will have p^{st} as its unique stationary distribution. Quite often, the approach to stationarity is faster with transition matrices that do not satisfy detailed balance, as discussed, e.g., in [Krau21].

8.4.2 Monte Carlo Algorithms in Statistical Mechanics

We now discuss how to build a Markov chain that enables us to sample thermodynamic ensembles. To be specific, we consider an Ising model defined on a d-dimensional cubic

lattice. The Hamiltonian is given by

$$\mathcal{H}(\sigma) = -\sum_{<i,j>} J\sigma_i\sigma_j - \sum_i h\sigma_i, \tag{8.78}$$

where $\sigma_i = \pm 1$, $i = 1, \ldots, N$, and the first sum is extended to nearest-neighbor pairs over the lattice. The collection $\sigma = (\sigma_i)$ of the spin values of all sites of the lattice defines the system's microstate. The thermodynamic observables (in the canonical ensemble) are obtained as averages of functions of σ with respect to the Boltzmann distribution

$$p_\sigma^{\text{eq}} = \frac{\mathrm{e}^{-\mathcal{H}(\sigma)/k_\text{B}T}}{Z}, \tag{8.79}$$

where Z is the partition function:

$$Z = \sum_\sigma \mathrm{e}^{-\mathcal{H}(\sigma)/k_\text{B}T}. \tag{8.80}$$

We thus obtain the expression of the internal energy,

$$E = \langle \mathcal{H} \rangle = \sum_\sigma \mathcal{H}(\sigma)\, p_\sigma^{\text{eq}}, \tag{8.81}$$

and that of the magnetization M;

$$M = \left\langle \sum_i \sigma_i \right\rangle = \sum_\sigma \left(\sum_i \sigma_i \right) p_\sigma^{\text{eq}}. \tag{8.82}$$

We want to define a Markov chain whose states correspond to the microstates σ of the system, and for which p^{eq} is the stationary distribution. If we impose that the transition matrix $\mathbb{W}$ possesses the detailed balance property, this result will be obtained if the matrix $\mathbb{W}$ is ergodic, and if for each pair of microstates (σ, σ') one has

$$W_{\sigma\sigma'} p_{\sigma'}^{\text{eq}} = W_{\sigma'\sigma} p_\sigma^{\text{eq}}. \tag{8.83}$$

This condition implies that either $W_{\sigma'\sigma} = 0$, i.e., the direct transition $\sigma' \longrightarrow \sigma$ is "forbidden" (but it has to take place sooner or later, following a finite number of steps; otherwise the matrix $\mathbb{W}$ is not ergodic) and the inverse transition $\sigma \longrightarrow \sigma'$ is forbidden as well, or the transition probabilities $W_{\sigma'\sigma}$ and $W_{\sigma\sigma'}$ can be expressed one as the function of the other:

$$W_{\sigma\sigma'} = W_{\sigma'\sigma} \frac{p_\sigma^{\text{eq}}}{p_{\sigma'}^{\text{eq}}} = W_{\sigma'\sigma} \exp\left[-\frac{\mathcal{H}(\sigma) - \mathcal{H}(\sigma')}{k_\text{B}T} \right]. \tag{8.84}$$

Let us remark that the partition function Z disappears from this relation, which is lucky for us, because if we knew how to calculate Z, we would have no need to resort to the simulation!

One therefore still has to choose the class of "allowed" transitions (for which $W_{\sigma\sigma'} > 0$) and define the $W_{\sigma\sigma'}$ so as to satisfy the relation (8.84), The inventiveness of researchers has been unleashed especially on solutions to the first problem, while the second has some standard solutions, the most common being the **Metropolis algorithm**. Given the state σ

of the system, one picks up a candidate state σ' such that $\sigma \longrightarrow \sigma'$ is an allowed transition. Then, the next state of the Markov chain is equal to σ' with probability

$$p(\sigma \longrightarrow \sigma') = \min\left(1, \exp\left\{-\left[\mathcal{H}(\sigma') - \mathcal{H}(\sigma)\right]/k_{\mathrm{B}}T\right\}\right); \tag{8.85}$$

otherwise, the next state is σ. In practice, if the transition goes downhill in energy, it takes place surely. Otherwise, it takes place only with probability $\exp(-\Delta\mathcal{H}/k_{\mathrm{B}}T)$, where $\Delta\mathcal{H} = \mathcal{H}(\sigma') - \mathcal{H}(\sigma)$. In this case, pick up a random number x, uniformly distributed between 0 and 1. Then, if $x < \exp(-\Delta\mathcal{H}/k_{\mathrm{B}}T)$, the transition takes place; otherwise, the system stays put. It is clear that with this choice, the condition (8.84) is satisfied.

The advantage of this algorithm is that it requires us to compare only the energy values of the two states connected by the transition, which is useful if the number of states accessible in one move is large.

While the delicate point of the algorithm is the choice of allowed transitions, the most computationally intensive one is the evaluation of $\Delta\mathcal{H}$. One usually chooses the allowed transitions so that $\Delta\mathcal{H}$ contains only a few terms—if, for example, σ and σ' differ only by our flipping the spin at i_0, all other spins remaining equal, one has

$$\Delta\mathcal{H} = 2\sigma_{i_0}\left(\sum_{j\in \mathrm{nn}(i_0)} J\sigma_j + h\right), \tag{8.86}$$

where the sum over j runs over the nearest neighbors of i_0. This expression requires the evaluation of a small number of terms, and above all, it can assume only a small number of different values. In order to save time, it is therefore useful to create a table with all possible values of $\exp(-\Delta\mathcal{H}/k_{\mathrm{B}}T)$ and use them when necessary.

In the case of the Ising model over a cubic lattice in any dimension d, with interactions between nearest neighbors, the algorithm's effectiveness can be improved by observing that the lattice can be divided into two sublattices: that in which the sum of the coordinates is an even multiple of the lattice constant a_0, and that in which it is an odd multiple. Now the nearest neighbors of a spin that belongs to the even sublattice all belong to the odd sublattice, and vice versa. It is therefore possible to update (decide whether to flip) all of a sublattice's spins without changing the value of $\Delta\mathcal{H}$ for the other spins in the sublattice. This operation can be carried out in parallel (if the machine allows for parallel computing), thus saving a great deal of time. If we update the two sublattices sequentially—first the even one (0) and then the odd one (1), and so forth—the transition matrix of the Markov chain is given by

$$\mathbb{W} = \mathbb{W}^{(1)}\mathbb{W}^{(0)}, \tag{8.87}$$

where $W^{(\alpha)}$ is the transition matrix for updating sublattice α ($\alpha = 0, 1$). Now, the $\mathbb{W}$ matrix does not have the detailed balance property, even if $\mathbb{W}^{(0)}$ and $\mathbb{W}^{(1)}$ have it.

Exercise 8.5 Prove this assertion, for example, by finding two configurations σ and σ' of a one-dimensional Ising model such that $W_{\sigma\sigma'} > 0$ while $W_{\sigma'\sigma} = 0$.

This, however, does not prevent the distribution from reaching equilibrium, since the $\mathbb{W}^{(\alpha)}$ separately admit p^{eq} as a stationary distribution, and the unicity of the stationary distribution is guaranteed by the ergodicity of $\mathbb{W}$.

Exercise 8.6 (Microcanonical ensemble) Describe a Markov chain (defined by the matrix $\mathbb{W}$ of the transition probabilities) that can sample the microcanonical ensemble:

$$p_\sigma^{\mathrm{micr}} = \begin{cases} 1/|\Gamma|, & \text{if } E < \mathcal{H}(\sigma) < E + \Delta E, \\ 0, & \text{otherwise,} \end{cases} \tag{8.88}$$

where $|\Gamma|$ is the number of accessible configurations.

Exercise 8.7 (One-dimensional Ising model) Consider a one-dimensional Ising model of N spins with free boundary conditions. Simulate the system with the Metropolis algorithm, alternately updating the even and odd spins. Evaluate the internal energy and the specific heat, using both the fluctuation-dissipation relation (3.93) $C = \operatorname{Var} \mathcal{H}/k_{\mathrm{B}}T^2$ and the relation $C \simeq \Delta \langle \mathcal{H} \rangle / \Delta T$. Compare with the exact result.

8.4.3 Random Sequences

> Anyone who attempts to generate random numbers by deterministic means is, of course, living in a state of sin.
>
> —J. von Neumann

The Monte Carlo method substitutes the "true" dynamics of the system being considered—a deterministic dynamics, in which the present microstate x of the system univocally dictates the next state x', and therefore the whole trajectory—by a fictitious "stochastic" dynamics, chosen to allow sampling the microstate distribution dictated by the relevant ensemble. The actual implementation of this dynamics requires obtaining from the computer very long sequences of random numbers, identically distributed and mutually independent. As mentioned by J. von Neumann in the quote, this looks perverse, even a conceptual impossibility.

It is a common occurrence, however, to use a computer to generate random numbers: fastidious souls introduce the distinction between *truly* random numbers, when they are generated by a natural process that we describe only probabilistically (like the number of disintegrations of some radioactive element) and the *pseudo*random (but reproducible) numbers that are produced by a computer. If one thinks about it carefully, however, the problem of why a deterministic system like a computer can produce sequences that are properly described as random is of the same nature as the basic problem of statistical mechanics—in other words, of why a deterministic system (described, e.g., by Newton's equations) can exhibit random behavior. A discussion of these problems would take us very far.

Here we adopt, following [Pres07, chap. 7], a pragmatic point of view: for our purposes, a program that produces sequences of random numbers is effective if, within the

application that uses these numbers, it produces statistically equivalent results to those obtained by using the sequences produced by another random number generator. Good random number generators, in other words, can be exchanged with one another within applications. This definition appears to be circular, because it is based on a comparison between different generators. It is useful, however, because it is a fact that good random number generators exist and have extensive fields of application. For a general discussion, see [Knut97, chap. 3].

More pragmatically, randomness, like beauty, is in the eye of the beholder. If one looks too closely, no sequence is sufficiently random—but the existence of hidden correlations does not necessarily make your results unusable. In any case, it is always a good idea to verify that the program you are using passes a series of tests aimed at revealing secret correlations or, if it does not pass them, to check that the relevant effects are not too important. A number of pitfalls are listed in [Pres07, p. 341], of which the most important is "never use the built-in generators in the C and C++ languages, especially `rand` and `srand`. These have no standard implementation and are often badly flawed." This advice is followed by this observation:

> If all scientific papers whose results are in doubt because of one of more of the above traps were to disappear from library shelves, there would be a gap on each shelf about as big as your fist. [Pres07, p. 342]

Before 1995, the most widely used random number generators were ultimately based on the **linear congruence** mechanism, which generates a sequence of integers $I_1, I_2, \ldots$ between 0 and $m-1$ (a large number) by means of the recurrence relation

$$I_{j+1} = aI_j + c \mod m. \tag{8.89}$$

Two integers, x and y, are congruent modulo m (which is written $x \equiv y \mod m$) if the remainders of their division by m are equal. Equation (8.89) therefore requires us to set I_{j+1} equal to the remainder of the division of $aI_j + c$ by m. By iterating the mapping, one obtains a sequence of integers contained between 0 and $m-1$. The linear congruence method has the advantage of being very fast, because it requires only very few operations for each call—this explains its wide use in earlier years. It was eventually realized, however, that it produces correlations between successive calls. If we take k successive random numbers to represent a point in k-dimensional space, then the points will arrange themselves on a few $(k-1)$–dimensional planes, instead of filling the space more or less uniformly. There will be at most $m^{1/k}$ planes, and usually there will be many fewer, unless a, c, and m are chosen very accurately. These correlations can (more or less) be reduced by reshuffling the numbers produced by the algorithm in the course of a certain number of successive calls.

Modern generators (after 1995) combine at least two (preferably unrelated) methods. Each method evolves independently of the others and does not share its state. The output should be obtained by a simple combination of their outputs that does not produce less random results than its operands. Quite often, the generators involve bitwise operators,

like the popular 64-bit Xorshift method due to G. Marsaglia [Mars03]. This is implemented by the pseudocode:

```
initialize:
    x (unsigned 64-bit, nonzero)

update:
    x = x ^ (x >> a1)
    x = x ^ (x << a2)
    x = x ^ (x >> a3)
```

Here `^` denotes the bitwise XOR operator, `>>` is the right bitwise shift, `<<` is the left bitwise shift, and `(a1,a2,a3)` are a triple of integers that must satisfy a number of constraints (a suitable triple is (11, 29, 14)). The iteration yields a sequence of unsigned integers between 1 and $2^{64} - 1$. The quality of the generator can be assessed by a number of statistical tests, like G. Marsaglia's Diehard suite [Mars95]. An updated and extended version, called Dieharder, is downloadable from R. G. Brown's General Tools page at Duke University [Brow20]. A number of methods for the generation of random numbers and their possible combinations are discussed in [Pres07, sec. 7.1].

8.4.4 Statistical Errors

Let us suppose that you have chosen your algorithm with great care, so that you have a $\mathbb{W}$ that ensures convergence toward the p^{eq} of interest. You choose an initial configuration $\sigma(0)$, and by applying your algorithm, you obtain a configuration sequence $(\sigma(0), \sigma(1), \ldots, \sigma(t), \ldots)$. You now wish to take advantage of this sequence to evaluate the observables of type $\langle A \rangle$, where $A(\sigma)$ is some function of the configuration σ, like the internal energy or the magnetization.

First, you will *discard* all configurations for a (long) initial stretch, $t < \mathcal{T}_0$. Your initial condition on the probability distribution is $p_\sigma(0) = \delta_{\sigma\sigma_0}$, and the probability distribution $p(t)$ will not be reasonably close to p^{eq} for quite some time (at least for the number of steps needed to allow us in principle to reach an arbitrary configuration σ starting from σ_0). The time $\mathcal{T}_0$ must be chosen by monitoring the values of the observables (averaged over relatively brief time intervals) until it becomes clear that they are varying only because of stochastic fluctuations, and no longer systematically.

But wait a minute: you can object that statistical mechanics is based on the fact that almost all the microstates that belong to the accessible region Γ have almost identical values of the thermodynamic observables, and therefore, for any $\sigma \in \Gamma$ and extensive observable A, one has

$$A(\sigma) \simeq \langle A \rangle \,, \tag{8.90}$$

up to subextensive corrections. Therefore, basically, all we would need to do is choose a random point in Γ and the entire lengthy simulation becomes useless.

Actually, it is not so easy to choose a random point. The accessible region is very large, on the order of e^{S/k_B}, but it is still a negligible portion of the entire configuration space. In

an Ising model in which $S = 0.5\, Nk_B \ln 2$, for example, one has

$$\frac{|\Gamma|}{2^N} \sim 2^{-0.5N}, \tag{8.91}$$

which, for $N = 100$, yields about 10^{-15}. It is therefore practically impossible for a point chosen at random to fall in Γ. The only way to move closer to Γ is to allow the system to do so spontaneously, in the course of its evolution.

Once we have discarded the initial stretch, we can set

$$\langle A \rangle \simeq [A]_{\mathcal{T}} = \frac{1}{\mathcal{T}} \sum_{t=\mathcal{T}_0}^{\mathcal{T}_0+\mathcal{T}} A(\sigma(t)), \tag{8.92}$$

where we have denoted by $[A]_{\mathcal{T}}$ the **running average** of A along a trajectory stretch of length $\mathcal{T}$. This expression yields $\langle A \rangle$ up to a statistical error $\Delta A_{\mathcal{T}}$, of the order of $\mathcal{T}^{-1/2}$. If the values of $A(\sigma(t))$ were independent of each other for different values of t, one would have

$$\left\langle \Delta A_{\mathcal{T}}^2 \right\rangle = \sum_{t=\mathcal{T}_0}^{\mathcal{T}_0+\mathcal{T}} \left(A(\sigma(t)) - \langle A \rangle\right)^2 \simeq \frac{1}{\mathcal{T}} \left\langle \left(A - \langle A \rangle\right)^2 \right\rangle. \tag{8.93}$$

The configurations obtained from the Markov chain, however, are not mutually independent: $\sigma(t)$ and $\sigma(t')$ will be independent only if $|t - t'|$ is larger than a characteristic correlation time τ_0. We can therefore expect

$$\left\langle \Delta A_{\mathcal{T}}^2 \right\rangle \simeq \frac{\tau_0}{\mathcal{T}} \left\langle \left(A - \langle A \rangle\right)^2 \right\rangle. \tag{8.94}$$

This can be much larger than the previous estimate. In order to be able to estimate the statistical error, it is necessary to have an idea of the size of τ_0. One way to do that is to monitor the **mutual overlap** of the configurations, defined by

$$q(t+\tau, t) = \frac{1}{N} \sum_i \sigma_i(t+\tau)\sigma_i(t). \tag{8.95}$$

If the configurations are independent, $q(t+\tau, t) \simeq m^2$, where $m = \langle \sigma_i \rangle$ is the magnetization per spin (at least in the ferromagnetic Ising model). Therefore, if we evaluate $q(t+\tau, t)$ as a function of τ, we can arrive at an estimate of τ_0 and hence also of the statistical error $\Delta A_{\mathcal{T}}$.

It is clearly very useful to choose one's algorithm in a way that makes τ_0 as small as possible. One can achieve this outcome by enlarging the space Γ_σ of configurations σ' accessible from a given configuration σ, making it so that the typical overlap between σ and $\sigma' \in \Gamma_\sigma$ is as small as possible. On the other hand, the transition $\sigma \longrightarrow \sigma'$ is in danger of not being accepted if $\mathcal{H}(\sigma')$ is much different from $\mathcal{H}(\sigma)$. Looking for very different configurations that have approximately the same energy is an interesting problem that has been solved using sophisticated methods in the case of the Ising model.

Let us note that $q(t+\tau, t)$ can reach a plateau as τ grows also if the system remains trapped in a metastable region of configuration space—in this case, all the quantities we will calculate will be relative to this metastable region and will have nothing to do with

equilibrium. This problem is important, because simulation times are always enormously shorter than the "real" times, and therefore a metastable state whose life span is much longer than the simulation time can have a very short life span in the real world.

8.4.5 Extrapolation to the Thermodynamic Limit

Let us now suppose that you have correctly evaluated the correlation time τ_0 and that you have obtained some trajectories long enough to produce averages $\langle A \rangle$ with a negligible statistical error. The problem is now to interpret the data you have obtained in this fashion, to obtain some information about the thermodynamics of your model. You might, for example, be interested in localizing the transition temperature T_c, estimating the critical exponents β and ν, and so on, for a two-dimensional Ising model. At this point (setting aside the problem of boundary conditions and other idealizations), you will have correctly simulated the behavior of a sample whose real-world size is equal to a few nanometers across, for a duration of some picoseconds. You will therefore want to extrapolate the data to obtain the behavior of the system in the thermodynamic limit.

The problem is that obviously the behavior of such a small system is qualitatively different from that of a macroscopic system. Let us consider the Ising model under T_c in a vanishing field. The infinite system exhibits a nonvanishing spontaneous magnetization (which is equal to $\pm m_0$ per spin), while we know that there cannot be spontaneous symmetry breaking in a finite system. In fact, if we evaluate the instantaneous magnetization $M(t)$ in the finite system

$$M(t) = \sum_i \sigma_i(t), \tag{8.96}$$

we will see that it normally fluctuates around $\pm N m_0 + \mathrm{O}\left(N^{1/2}\right)$, but its sign changes from time to time. If we attempted to calculate $\langle M \rangle$ by averaging $M(t)$ over a time interval in which these inversions do not appear, we would not reduce enough our statistical errors. It is therefore necessary to calculate some quantities that will not be influenced by these inversions. In our case, it is convenient to introduce the quantity m_{RMS}, defined by

$$m_{\mathrm{RMS}}^2 = \frac{1}{N^2} \left\langle M^2(t) \right\rangle = \frac{1}{N^2} \left\langle \left[\sum_i \sigma_i(t) \right]^2 \right\rangle. \tag{8.97}$$

This quantity is well defined but has the drawback that it also remains different from zero for $T > T_c$. Therefore, we cannot immediately determine T_c in the usual way, as the temperature above which the spontaneous magnetization vanishes. This property of m_{RMS} is, however, actually an advantage, because it allows us to reduce the systematic error over m_{RMS} to values smaller than m_{RMS} itself (something that would be rather difficult to do for quantities that vanish) by performing long enough simulation runs, and therefore allows us to produce reliable m_{RMS} curves as a function of T.

In order to exploit these data, one makes use of the theory of **finite-size scaling**, which I now attempt to summarize. Let us consider a sample of linear dimensions L and discuss its spin susceptibility $\chi(L, T)$, for example. On the one hand, we expect $\chi(L, T)$ to have the

following scaling expression:

$$\chi(L, T) = L^{\gamma}\hat{\chi}\left(L/\xi(T)\right), \tag{8.98}$$

where $\xi(T) \propto |T - T_c|^{-\nu}$ is the coherence length and γ is an exponent that we now determine. On the one hand, if the coherence length ξ is much smaller than L, we expect the sample to behave substantially like an infinite system. Therefore, for $L \gg \xi$, one should have $\chi \propto |T - T_c|^{-\gamma}$, i.e., $\chi \propto \xi^{\gamma/\nu}$. Therefore, one must have $\hat{\chi}(x) \propto x^{-\gamma/\nu}$ for $x \gg 1$ and, in order for χ to be independent of L, $\gamma = \gamma/\nu$. On the other hand, for $L \ll \xi(T)$, we expect χ (and therefore $\hat{\chi}(L/\xi)$) to be independent of ξ, and therefore $\chi(L, \xi) \propto L^{\gamma/\nu}$.

Let us now observe that if $T > T_c$,

$$m^2_{\mathrm{RMS}} = \frac{1}{N^2}\sum_{ij}\sigma_i\sigma_j = \frac{k_B T}{N}\chi(L, T). \tag{8.99}$$

Substituting the above results for $\chi(L, T)$, we obtain

$$m^2_{\mathrm{RMS}} = L^{-d+\gamma/\nu}\hat{\chi}\left(\frac{L}{\xi}\right) = L^{-2\beta/\nu}\hat{\chi}\left(L|T - T_c|^{\nu}\right), \tag{8.100}$$

where we have exploited the scaling law

$$\beta = \frac{1}{2}(d\nu - \gamma). \tag{8.101}$$

A similar reasoning yields, for $T < T_c$, $m_{\mathrm{RMS}}(L, T) = L^{\beta/\nu}\hat{m}(L|T - T_c|^{\nu})$, where $\hat{m}(x) \sim x^{\beta/\nu}$ when $x \gg 1$. Thus, plotting $m^2_{\mathrm{RMS}}L^{2\beta/\nu}$ as a function of $x = L|T - T_c|^{\nu}$ for various values of L, all data should fall on a single line. This requires us of course to know the values of T_c and of the exponents β and ν. Below T_c, one has $m^2_{\mathrm{MRS}} \simeq L^{-2\beta/\nu}$

It is possible to determine these values iteratively by observing that, in view of eq. (8.100), the temperature $T_0(L)$, where $\mathrm{d}^2 m_{\mathrm{RMS}}/\mathrm{d}T^2 = 0$, behaves as

$$T_0(L) = T_c + \mathrm{O}\left(L^{-1/\nu}\right), \tag{8.102}$$

while $m_{\mathrm{RMS}}\left(T_0(L)\right)$ scales as $L^{-\beta/\nu}$. Thus, by evaluating $T_0(L)$—in other words, the values of T where m_{RMS} exhibits an inflection point—one can determine T_c, ν and β/ν. The quality of the result can then be checked by replotting all the data in the form dictated by eq. (8.100). An example of this *data collapse* for the two-dimensional Ising model (where $\beta = 1/8$, $\gamma = 7/4$, and $\nu = 1$) is shown in figure 8.1.

A different approach focuses on the distribution $p_L(m)$ of the instantaneous values of the magnetization per spin $m = M/N$. For $T > T_c$ this distribution is a Gaussian, with a width determined by the susceptibility per spin $\chi(L, T)$:

$$p_L(m) \propto \exp\left[-\frac{m^2}{2k_B T\,\chi(L, T)}\right]. \tag{8.103}$$

For $T < T_c$, the distribution exhibits two peaks, and, at low temperatures, becomes the sum of two narrow Gaussians. For T close to T_c, we expect it to have a scaling form:

$$p_L(m) = L^{\gamma}\hat{p}\left(mL^{\gamma}, L/\xi(T)\right). \tag{8.104}$$

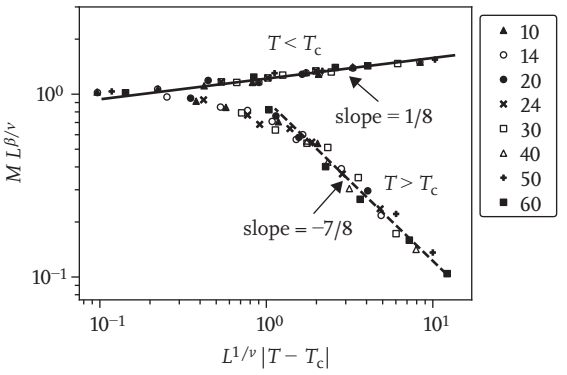


FIGURE 8.1. Illustration of finite-size scaling. Rescaled magnetization $ML^{\beta/\nu}$ for the two-dimensional Ising model, as a function of the rescaled distance from the critical temperature $L^{1/\nu}\,|T - T_c|^{\nu}$, for both $T > T_c$ and $T < T_c$. The legend lists the values of L. Adapted from [Milc86].

The exponent γ can be determined by looking at the average of $|m|$:

$$\begin{aligned}\langle |m| \rangle &= L^{\gamma} \int \mathrm{d}m\ |m|\ \hat{p}(mL^{\gamma}, L/\xi(T)) = L^{-\gamma} \int \mathrm{d}z\ |z|\ \hat{p}\,(z, L/\xi(T)) \\ &= L^{-\gamma}\Phi(L/\xi(T)).\end{aligned} \tag{8.105}$$

This is the usual scaling form of the order parameter, and it implies $\gamma = \beta/\nu$. Let us now consider the moments $\langle m^2 \rangle$ and $\langle m^4 \rangle$. We have

$$\langle m^2 \rangle = L^{-2/\gamma} \int \mathrm{d}z\, z^2 \hat{p}\,(z, L/\xi(T))\,, \tag{8.106}$$

$$\langle m^4 \rangle = L^{-4/\gamma} \int \mathrm{d}x\, z^4\, \hat{p}\,(z, L/\xi(T))\,. \tag{8.107}$$

Therefore, the ratio

$$R = \frac{\langle m^4 \rangle}{\langle m^2 \rangle^2} = \Psi(L/\xi(T)) \tag{8.108}$$

approaches a well-defined limit (independent of L) for $\xi(T) \to \infty$. On the other hand, for $T \gg T_c$, $R = 3$, while we expect $R \to 1$ for $T \to 0$. Therefore, plotting R vs. T for different values of L, the different curves will intersect at a point with a value between 1 and 3, in correspondence of the critical temperature for the infinite system.

8.5 Umbrella Sampling

Computing time is not free. Say that you have finally obtained a long Monte Carlo run for your Ising model, at the temperature T_0 and magnetic field h. It would not be a good idea to throw your precious data away before extracting as much relevant information as

possible. For example, you could exploit it so as to obtain data for different values of the temperature and the magnetic field.

Indeed, if the sequence $(\sigma(0), \sigma(1), \ldots)$ samples the canonical distribution at the temperature T, the average of the observable A at the temperature T can be evaluated by use of the relation

$$\begin{aligned}
\langle A \rangle_T &= \frac{1}{Z(T)} \sum_\sigma A(\sigma)\, \mathrm{e}^{-\mathcal{H}(\sigma)/k_\mathrm{B}T} \\
&= \frac{1}{Z(T_0)} \frac{Z(T_0)}{Z(T)} \sum_\sigma A(\sigma)\, \exp\left[-\mathcal{H}(\sigma)\left(\frac{1}{k_\mathrm{B}T} - \frac{1}{k_\mathrm{B}T_0}\right)\right] \mathrm{e}^{-\mathcal{H}(\sigma)/k_\mathrm{B}T_0} \\
&= \left\langle A \exp\left[-\mathcal{H} \cdot \left(\frac{1}{k_\mathrm{B}T} - \frac{1}{k_\mathrm{B}T_0}\right)\right]\right\rangle_{T_0} \\
&\quad \times \left\langle \exp\left[-\mathcal{H} \cdot \left(\frac{1}{k_\mathrm{B}T} - \frac{1}{k_\mathrm{B}T_0}\right)\right]\right\rangle_{T_0}^{-1},
\end{aligned} \tag{8.109}$$

where $\langle \ldots \rangle_T$ denotes the canonical average at the temperature T.

Thus the average at temperature T can be obtained by evaluating the means of suitable variables along trajectories sampled at the different temperature T_0. One can calculate the averages relative to the different values of the magnetic field (and so on) in a similar fashion. One would do well, however, to keep in mind that this method works only if the perturbation (the argument of the exponential in the preceding expression) is very small. If the exponential factor becomes very large or very small, the results will no longer be trustworthy. Practically speaking, it is convenient to *always* use this trick to evaluate the effect of a small variation in the parameters, thus obtaining, e.g., the susceptibility or the specific heat. One can then compare these susceptibilities with those predicted from the estimated fluctuations by the fluctuation-dissipation theorem. If the results do not agree, one can infer that something is not working in the simulation—for example, that equilibrium has not actually been reached, or that the correlation time has been underestimated.

In a similar fashion, we can sample the canonical distribution with a Hamiltonian $\mathcal{H}_0$, which is not identical to the one (let us call it $\mathcal{H}$) that we are interested in, as long as we are careful to take the correction into account. By denoting the average over the Hamiltonian $\mathcal{H}$ by $\langle \ldots \rangle_\mathcal{H}$, we have in general

$$\langle A \rangle_\mathcal{H} = \frac{\left\langle A\, \mathrm{e}^{-(\mathcal{H}-\mathcal{H}_0)/k_\mathrm{B}T} \right\rangle_{\mathcal{H}_0}}{\left\langle \mathrm{e}^{-(\mathcal{H}-\mathcal{H}_0)/k_\mathrm{B}T} \right\rangle_{\mathcal{H}_0}}. \tag{8.110}$$

This technique is called **umbrella sampling**. In addition to being used for the study of perturbations with respect to a given Hamiltonian, it can be used for the study of rare events. We can, in fact, be interested in the properties of a rare class of configurations in which certain properties are satisfied, and we can introduce $\mathcal{H}_0$ in order to facilitate their appearance.

A fairly frequent case is that when one wants to study the probability distribution of the order parameter in finite samples. This is the same as studying the $p_L(m)$ distribution we discussed in the preceding section using a histogram. The shape of the histogram actually allows us to distinguish between continuous and discontinuous transitions. In the vicinity of a discontinuous transition, $p_L(m)$ exhibits peaks close to the different equilibrium values of m, while in the case of a continuous transition, the distribution is much broader. On the other hand, during a run at temperature $T < T_c$, only those configurations in which m is not far from the equilibrium value m_0 will be "spontaneously" sampled. So we can introduce a perturbation $\mathcal{V}(\sigma)$ by defining $\mathcal{H}^* = \mathcal{H} + \mathcal{V}$, choosing $\mathcal{V}$ such that $\mathcal{H}^*$ favors values of m different from m_0. For example, we can set $\mathcal{V}(\sigma) = 0$ if m belongs to a window $m^* \leq m \leq m^* + \Delta m$, and $V(\sigma) = +\infty$ otherwise. We can thus evaluate the histogram $p^*(m)$ in the window. We can then collate the histogram by imposing the continuity of $\ln p^*(m)$ in adjacent windows. It is better to look at $\ln p^*(m)$ than at $p^*(m)$ itself, since it varies much more slowly, and has a physical interpretation in term of availability. In this way, it is possible to reconstruct the behavior of the entire $p_L(m)$.

Exercise 8.8 Evaluate the specific heat in the one-dimensional Ising model using umbrella sampling and check the validity of the fluctuation-dissipation relation.

Exercise 8.9 (Widom insertion method) The chemical potential of a fluid can be evaluated via a molecular dynamics or Monte Carlo simulation by the following method, due to B. Widom [Wido63].

1. Express the chemical potential μ of a simple fluid with N particles in terms of the partition functions Z_{N+1} and Z_N of a system with respectively $N+1$ and N particles. Deduce its expression in terms of the configuration integrals Ω_{N+1} and Ω_N.

2. Express Ω_{N+1} as the average of the Boltzmann factor of a "ghost particle" inserted at random positions within the system, expressed in terms of its interaction with the N remaining particles. Deduce the expression of μ.

8.6 Accelerated Monte Carlo Methods

In this section I briefly report on two techniques based on the Monte Carlo procedure, which have been introduced to deal with systems with complex energy landscapes. Such systems are characterized by sluggish dynamics, which prevents them from reaching equilibrium with a simple Monte Carlo simulation at a fixed temperature.

The first method has been successfully applied to the search for optima of systems with a large configuration space. Consider a system with discrete configuration states x, and an associated cost function $\mathcal{H}(x)$. Then, looking at the states that minimize the cost function is akin to looking for the minimal-energy states of a system with the same space of configurations and with Hamiltonian $\mathcal{H}(x)$. If the system is at equilibrium at a very low

temperature, its probability distribution (proportional to the Boltzmann factor $e^{-\mathcal{H}(x)/k_B T}$) will be concentrated around the low-lying minima of $\mathcal{H}(x)$.

The catch is that when the system is such that there are a great number of states with low values of the energy, it is rather difficult to let the system equilibrate at low temperature by a direct Monte Carlo procedure. Indeed, in this procedure we need to start from a configuration picked at random, that is, in the absence of detailed information about the system. By performing the Monte Carlo procedure at low temperature, the system will rapidly lose energy, and it is highly probable that it will quickly get stuck in a local minimum, surrounded by states of high energy. Then the Metropolis algorithm will assign a practically vanishing probability to a move that brings the system away from the local minimum. Therefore, a direct implementation of the Monte Carlo process at low temperature will not allow us to bring the system to the actual low-lying minima.

Simulated annealing is a Monte Carlo simulation protocol that attacks this kind of problems. The method simulates the heating and slow cooling of metals in a procedure aimed at removing crystal defects. This can be obtained by performing a Monte Carlo simulation in which the temperature varies according to an annealing schedule, starting from a high (or even infinite) value, and slowly decreasing toward the desired value. The cooling speed must be chosen to be neither too fast, to allow at least a partial equilibration at the current temperature, nor too slow, to avoid useless waste of computation time. The technique was introduced by several researchers, in particular by M. Pincus [Pinc70], and received its name by S. Kirkpatrick and collaborators [Kirk83]. An accessible discussion of the method can be found in the textbook by W. H. Press and collaborators [Pres07, 10.12].

To be specific, let us consider a classic combinatorial optimization problem: the problem of the travelling salesman. The salesman must visit N cities, one by one, starting and ending in a given city. Assume, for example, that the trip between any two cities has a cost proportional to the distance between them. The problem is minimizing the cost of the trip by choosing a suitable order in which to visit the cities. The difficulty is that the number of possible orderings increases very fast with the number N of cities: for $N = 15$ there are more than 1,300 billion possible orderings. Therefore a brute force exploration is out of the question. It can actually be shown that this problem belongs to the class of NP-hard problems, namely, it is a problem whose solution can be checked in time that does not increase too fast with the size of the problem (that is, the checking time increases "polynomially," as a power of N) and one that can be used to simulate every other problem having the same property. In this sense, once this problem is solved, one can find in principle (in polynomial time!) the solution of any other problem of its kind. It is not known if a solution to this kind of problem can be found in polynomial time: this is known as the "P=NP" problem.

The simulated annealing protocol runs as follows. Define the states x as the permutations $x = (x_1, x_2, \ldots, x_N)$ of the N cities, representing a possible order of visits, and assign to x the total length $\mathcal{H}(x)$ of the corresponding trip:

$$\mathcal{H}(x) = \sum_{k=1}^{N} d(x_k, x_{k+1}), \tag{8.111}$$

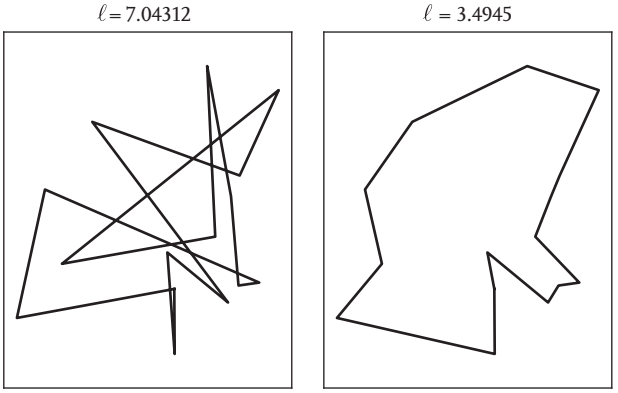


FIGURE 8.2. Example of the solution of a travelling salesman problem by simulated annealing. The $N = 15$ cities are placed at random in the unit square. Left: Initial path. Right: Solution found by simulated annealing with $\lambda = 100$ and $\mu = 0.9$. The path length ℓ is also shown.

where $d(x, x')$ is the distance between cities x and x' $(x, x' \in \{1, 2, \ldots, N\})$, and $x_{N+1} = x_1$, since the salesman must end the trip in the city he started from. Pick up one permutation $x(0)$ at random and set the temperature T to a moderately large value $T(0)$.

Then perform the following procedure:

1. Perform a Monte Carlo simulation at the temperature T for a number of steps given by λN, where λ is a suitably large number (say 100). At each step, the system chooses at random between two moves:

   Reversal: Pick up two cities at random, and inverse the path between them.

   Transport: Pick up two cities at random and one that does not lie in the path between them. Then remove the path between the first two and paste it between the third city and the one that immediately follows it.

   The move is accepted or rejected according to the usual Metropolis protocol, namely, it is accepted with probability 1 if it reduces the cost, or with probability $e^{-\Delta\mathcal{H}/k_B T}$ if the cost $\mathcal{H}$ increases by $\Delta\mathcal{H} > 0$. In this way the visiting order is modified.

2. Reduce the temperature T by a factor μ close to, but smaller than, 1 (for example, 0.9), and go back to step 1.

In this way, the system is gently led to exploring states with lower and lower values of $\mathcal{H}(x)$. The procedure can end when in a Monte Carlo run of λN attempts there are no successful moves, meaning that there are no more accessible low-lying states. The result of such a protocol on a model of $N = 15$ cities on a plane is shown in fig. 8.2.

This method is well suited for finding optimal solutions to combinatorial optimization problems, and it has been cleverly extended to problems with continuous phase space. On the other hand, it is not well adapted to exploring the behavior of complex systems at low temperatures. Indeed, by its very nature, simulated annealing brings the system out of

equilibrium. Moreover, one might be interested in the behavior of the system at different temperatures, not necessarily low.

This problem may be tackled by a different procedure, called **parallel tempering**. This is a computation-intensive procedure that involves simulating in parallel several identical copies of the system, each at a different temperature. (Thus the procedure is well adapted to implementation in computer clusters, with each CPU dedicated to a different replica.) From time to time, replicas at adjacent temperatures exchange their states. More exactly, the exchange is proposed and accepted or rejected according to a variation of the Metropolis rule. Given the energies (E_1, E_2) of the two replicas involved in the swap and their respective temperatures (T_1, T_2), the acceptance probability p is given by

$$p = \min\left(1, \exp\left[(E_1 - E_2)\left(\frac{1}{k_B T_1} - \frac{1}{k_B T_2}\right)\right]\right). \tag{8.112}$$

In this way the detailed-balance condition is satisfied, and the different replicas eventually sample the equilibrium distribution at the given temperature in a correct way. This method was introduced by R. Swendsen and J.-S. Wang [Swen86], and later improved by several researchers, among them E. Marinari and G. Parisi [Mari92].

To optimize the performance of the method, the temperatures assigned to the replicas must be chosen with care. A necessary condition is that the energy values of neighboring replicas overlap to some extent, so that the success probability of an exchange is neither 0 nor 1 (it appears [Rath05] that the best performance is reached for an acceptance ratio of about 20%). To avoid waste of computational time, it is best to avoid substantial overlap of the energy values with next-nearest neighboring replicas. If the distribution of energy values is Gaussian, it has been shown [Kofk02] that the average acceptance ratio of the swap is given by

$$\langle A \rangle = \operatorname{erfc}\left[\left(\frac{C_V}{2}\right)^{1/2} \frac{1 - T_1/T_2}{\sqrt{1 + (T_1/T_2)^2}}\right], \tag{8.113}$$

where C_V is the heat capacity at constant volume of the system, assumed to be constant between T_1 and T_2. Thus, if C_V a constant, a geometric progression of temperatures (that is, $T_k/T_{k+1} = \text{const.}$) results in a uniform average acceptance ratio. The procedure has been extended to systems simulated by molecular dynamics. Moreover, sophisticated method have been devised to optimize the performance of the procedure. See, for example, the review by D. J. Earl and M. W. Deem [Earl05].

8.7 Discussion

One of the most interesting aspects of numerical simulation is the way in which it allows one to clarify the fundamental conceptual problems of statistical mechanics. We have actually seen how a suitably defined Markov chain allows us to sample the canonical ensemble, for example. The runs one can actually perform, however, allow us to evaluate means

over only a ridiculously small number of configurations. For $N = 100$, for instance, we have about 10^{30} possible configurations—in other words, ten million trillion trillions. How many truly independent configurations will you manage to sample? If you are patient, you can reach several hundred thousand (for 100 spins). You are therefore not sampling the Boltzmann distribution in the same sense in which, for example, you can check if a die is fair by throwing it 10,000 times and checking that the frequencies of the various results are equal to 1/6 within 1%.

Indeed, we are not aiming to sample the distribution in phase space, but only to evaluate the averages of observables. Let us consider the correlation function $\langle \sigma_i \sigma_j \rangle$. It depends on the four possible configurations of σ_i and σ_j. Thus, if we have 100,000 independent configurations, we can be fairly confident about our results, since the statistical errors will be small. On the other hand, it will be difficult to evaluate the correlations for ℓ spins, where ℓ is only slightly larger (6 or 8), because the number of configurations they can exhibit begins to become too large. More generally, we will have to be careful not to squeeze our data too much—in other words, not to base our conclusions on events that, in our run, occurred a small number of times. If we really have to, we can use umbrella sampling to obtain a statistically significant sample.

Lastly, let us remark that we were able to obtain only averages of *observables*. There are important thermodynamic quantities that are not observable: first entropy, then free energy, and so on. Strictly speaking, one could express the partition function (and therefore free energy) by an average. Since

$$\left\langle \mathrm{e}^{\mathcal{H}/k_\mathrm{B}T} \right\rangle = \frac{1}{Z} \sum_\sigma \mathrm{e}^{\mathcal{H}(\sigma)/k_\mathrm{B}T} \mathrm{e}^{-\mathcal{H}(\sigma)/k_\mathrm{B}T} = \frac{2^N}{Z}, \tag{8.114}$$

we can write

$$Z = \frac{2^N}{\left\langle \mathrm{e}^{\mathcal{H}/k_\mathrm{B}T} \right\rangle}. \tag{8.115}$$

Since $\mathcal{H}$ is extensive, however, the exponential will be too large in a too extended region of configuration space for our results to be reliable. We had therefore better look for some methods to directly calculate better-behaved quantities, for instance, the entropy S. One of the most frequently used methods (which works if the algorithm reaches equilibrium very fast) is to integrate the specific heat, as discussed in sec. 8.2.1. There are also techniques that allow one to evaluate S directly, such as the *coincidence method*, due to S.-K. Ma [Ma85, chap. 25].

Recommended Reading

Molecular dynamics is discussed by M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids*, Oxford: Oxford University Press, 1987; by D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Academic Press, 2001; and several others. A popular-science-level book on the subject, also containing a discussion of its

history, is due to G. Battimelli, G. Ciccotti, and P. Greco, *Computer Meets Theoretical Physics: The New Frontier of Molecular Simulation*, Berlin: Springer, 2020. An introductory text to the Monte Carlo method is K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics*, Berlin: Springer, 1988. More advanced contributions are found in K. Binder, *Monte Carlo Methods in Statistical Mechanics*, (6th ed.) Berlin: Springer, 2019. A reference that also ventures into more general subjects is W. Krauth, *Statistical Mechanics: Algorithms and Computations*, Oxford: Oxford University Press, 2006. L. Böttcher and H. J. Herrmann, *Computational Statistical Physics*, Cambridge UK: Cambridge University Press, 2021, contains a comprehensive and didactic introduction to the subject. Nontraditional applications of numerical simulation are discussed in S. Moss de Oliveira, P. M. C. de Oliveira, and D. Stauffer, *Evolution, Money, War, and Computers*, Stuttgart: Teubner, 1999.

9 Dynamics

> Theories are made to seduce aesthetes, irritate philistines, and make everyone else laugh.
>
> —Amélie Nothomb

The time evolution of a system in statistical mechanics is described either by the classical Liouville equation or by its quantum analog. These equations deal with the evolution of the probability distribution function or the density matrix, at the level of microscopic states. For a macroscopic system, this description level is so detailed as to be unmanageable. Nevertheless, since these equations are compatible with the equilibrium properties described by equilibrium statistical mechanics, it is possible to identify a number of quite general and experimentally relevant features of the dynamic behavior of observables. It is hard and still essentially an open problem to derive these features directly from the microscopic equations of motion. However, it is possible to treat these features at a mesoscopic level, that is, at an intermediate level between the microscopic one (corresponding, for example, to the $6N$ degrees of freedom of a classical fluid) and the macroscopic one (corresponding to the thermodynamic degrees of freedom). The paradigm of this description is the irregular motion of a comparatively large, but still microscopic, particle immersed in a solvent at thermal equilibrium, known as Brownian motion.

9.1 Brownian Motion

Brownian motion was "officially" discovered by the botanist R. Brown in 1827,[1] but remained a puzzle until the beginning of the twentieth century. J. Perrin, in his classic book *Atoms*, describes it in the following manner [Perr16, p. 83–85]:

> We have only to examine under the microscope a collection of small particles suspended in water to notice at once that each of them, instead of sinking steadily, is quickened by an extremely lively and wholly haphazard movement. Each particle spins hither and thither, rises, sinks, rises again, without ever tending to come to rest. . . . Of course, the phenomenon is not

[1] The phenomenon had been known informally for a while, and had been described, for example, by J. Ingenhousz in 1785.

> confined to suspensions in water, but takes place in all fluids, though more actively the less viscous the fluid. Thus it is just perceptible in glycerin, and extremely active, on the other hand, in gases. . . . In a given fluid the size of the grains is of great importance, the agitation being the more active the smaller the grains. This property was pointed out by Brown at the time of his original discovery. The nature of the grains appears to exert little influence, if any at all.

Already in the nineteenth century it was suspected, in particular by C. Wiener in 1863 and by L. G. Gouy in 1888, that the motion of the observed particle was due to its collisions with the fluid molecules, but this hunch did not lead to a quantitative understanding of the phenomenon. Einstein [Eins05], in one his fundamental works of 1905, and M. Smoluchowski [Smol06] were the first to provide a quantitative explanation of Brownian motion. The subject was taken up again by Einstein, by P. Langevin, and by L. S. Ornstein, and experimentally verified by J. Perrin, T. Svedberg, and others at the beginning of the last century. It represents one of the fundamental proofs of the corpuscular nature of matter. The most important papers by Einstein on this subject are collected in [Eins56]. Other important papers are [Lang08], [Orns20], [Perr08], and [Sved06].

Let us consider a particle immersed in a homogeneous solvent at thermal equilibrium, at a temperature T. Let the size and the mass of the particle be much larger than the corresponding values for the solvent molecules. If the particle moves with a velocity $\boldsymbol{v}$, because of resistance of the fluid, it is subject to a force $\boldsymbol{F} = -\gamma \boldsymbol{v}$, where $\gamma > 0$ is the **damping coefficient** that depends both on the geometry of the particle and on the viscosity of the fluid, and we neglect higher-order terms in v. For a spherical particle of radius R, immersed in a fluid of viscosity η, Stokes's law gives

$$\gamma = 6\pi\eta R. \tag{9.1}$$

On the other hand, because of the thermal agitation animating the molecules of the liquid, the particle will be subject to a random force $\boldsymbol{f}(t) = (f_1(t), f_2(t), f_3(t))$. We can reasonably assume that this random force has the following properties:

Isotropy: The force does not favor any particular direction. (For simplicity's sake, let us momentarily disregard the effects of external forces, such as weight.)

Noncorrelation: The force fluctuates at each instant, and its value at an instant t is independent of that at instant t', as soon as $|t - t'|$ is larger than a microscopic time τ_{micr}.

Gaussian distribution: The force $\boldsymbol{f}(t)$ is the result of the independent actions performed at every instant by a large number of molecules. As a result of the central limit theorem (cf. sec. A.6), we expect that its distribution is close to a Gaussian.

We shall now evaluate the dynamics of the particle, and we impose that it satisfy the conditions imposed by equilibrium statistical mechanics, since the Brownian particle differs from the molecules of the solvent *only by its size*.

Since the distribution of the random force is Gaussian, it is determined by its average and its correlation, which can be derived from the other two properties. We thus

obtain

$$\langle f_\alpha(t)\rangle = 0, \qquad \alpha = 1,2,3, \quad \forall t; \tag{9.2}$$

$$\langle f_\alpha(t) f_\beta(t')\rangle = \Lambda\, \delta_{\alpha\beta}\, \Delta(|t-t'|), \qquad \alpha,\beta = 1,2,3, \quad \forall t,t'. \tag{9.3}$$

In eq. (9.3), $\Delta(t)$ is a function that rapidly tends to zero when its argument is larger than the very short time τ_{micr}, and it is arbitrarily normalized so that $\int_{-\infty}^{+\infty} \mathrm{d}t\, \Delta(t) = 1$. It can therefore by represented by a Dirac delta, as long as we don't investigate the properties of the motion on a timescale shorter than τ_{micr}.

Einstein's idea is that, on the one hand, the motion of the Brownian particle is due to the random collisions it has with the fluid molecules, and, on the other hand, since it differs from the fluid molecules only in its size, it must satisfy the requirements of statistical mechanics. Here we follow this idea by evaluating the average velocity square of the particle and ask that it satisfy the equipartition theorem.

By applying Newton's law to a particle of mass m, we obtain the **Langevin equation**:

$$m\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -\gamma\, \boldsymbol{v} + \boldsymbol{f}(t). \tag{9.4}$$

This is an example of a **stochastic differential equation**—in other words, of a differential equation in whose definition random quantities appear, such as the random force $\boldsymbol{f}(t)$. We have derived it (and will treat it) in a heuristic manner, and the notation we use does not with stand mathematical scrutiny: it is, however, possible to give it the rigorous definition and treatment due, among others, to N. Wiener [Wien30] and K. Ito [Ito44]. An accessible introduction to the theory and literature of stochastic differential equations is available in [Gard09] and [Guer81].

To simplify the discussion it is convenient to consider the problem in one dimension for the time being. By denoting the only components of $\boldsymbol{v}$ and $\boldsymbol{f}$ as v and f respectively, we obtain

$$m\frac{\mathrm{d}v}{\mathrm{d}t} = -\gamma\, v + f(t). \tag{9.5}$$

This is a linear equation in $v(t)$. It can be formally integrated, and one obtains the solution

$$v(t) = v(0)\, \exp\left(-\frac{\gamma}{m}t\right) + \frac{1}{m}\int_0^t \mathrm{d}t'\, \exp\left[-\frac{\gamma}{m}(t-t')\right] f(t'), \tag{9.6}$$

where $v(0)$ is the value of $v(t)$ at $t = 0$. We see that $v(t)$ is a linear functional of $f(t)$. Since $f(t)$ is Gaussian, and since linear functions of random Gaussian variables are Gaussian, $v(t)$ (for a given value of $v(0)$) is also a random Gaussian process. By taking the average of both sides of eq. (9.6), and keeping in mind eq. (9.2), we obtain

$$\begin{aligned}\langle v^2(t)\rangle = \exp\left(-\frac{2\gamma}{m}t\right)\langle v^2(0)\rangle + \frac{2}{m}\int_0^t \mathrm{d}t'\, \exp\left[-\frac{\gamma}{m}(t-t')\right]\langle v(0) f(t')\rangle \\ + \frac{1}{m^2}\int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t''\, \exp\left\{-\frac{\gamma}{m}\left[2t - (t'+t'')\right]\right\}\langle f(t') f(t'')\rangle .\end{aligned} \tag{9.7}$$

The second term vanishes because $v(0)$ and $f(t)$ are independent, and one has $\langle f(t)\rangle = 0$, $\forall t$. The third term contributes only for $t' = t''$ because of eq. (9.3). We thus obtain

$$\begin{aligned}\langle v^2(t)\rangle &= \exp\left(-\frac{2\gamma}{m}t\right)\langle v^2(0)\rangle + \frac{\Lambda}{m^2}\int_0^t \mathrm{d}t'\ \exp\left[-\frac{2\gamma}{m}(t-t')\right] \\ &= \exp\left(-\frac{2\gamma}{m}t\right)\langle v^2(0)\rangle + \frac{\Lambda}{2m\gamma}\left[1-\exp\left(-\frac{2\gamma}{m}t\right)\right].\end{aligned} \tag{9.8}$$

We see that after a short time (of the order of $(m/2\gamma)$) one has $\langle v^2(t)\rangle \simeq \Lambda/2m\gamma$.

We recall that the solvent and the particle are at equilibrium at the temperature T. This implies, in particular, that the particle must satisfy the equipartition theorem, which stipulates that

$$\frac{1}{2}m\langle v^2\rangle = \frac{1}{2}k_{\mathrm{B}}T. \tag{9.9}$$

Comparing this relation with eq. (9.8), we obtain

$$\Lambda = 2\gamma\, k_{\mathrm{B}}T. \tag{9.10}$$

Notice that if eq. (9.9) is satisfied at $t = 0$, eq. (9.8) implies that it remains satisfied at all subsequent times. Equation (9.10) exhibits a relation between γ, which describes the mobility of the Brownian particle, and Λ, which expresses the fluctuations of the random force $f(t)$. This relation has the drawback that Λ is not directly measurable.

9.2 Diffusion Coefficient and the Einstein Relation

It is possible, however, to obtain a relation between γ and a directly observable quantity connected with the amplitude of the excursions of the Brownian particle. Following Langevin [Lang08], let us multiply both sides of eq. (9.4) by $x(t)$ and define

$$\Delta x(t) = x(t) - x(0), \tag{9.11}$$

$$z(t) = \frac{\mathrm{d}}{\mathrm{d}t}\Delta x^2(t). \tag{9.12}$$

We obtain

$$\frac{1}{2}m\frac{\mathrm{d}z}{\mathrm{d}t} - mv^2 = -\frac{\gamma}{2}z + x(t)f(t). \tag{9.13}$$

Let us take the average of this relation, keeping in mind the equipartition theorem, as well as the fact that $f(t)$ is independent of $x(t)$ and has a vanishing average. We obtain

$$\frac{1}{2}m\frac{\mathrm{d}\langle z\rangle}{\mathrm{d}t} + \frac{\gamma}{2}\langle z\rangle = k_{\mathrm{B}}T. \tag{9.14}$$

Therefore

$$\langle z(t)\rangle = \frac{2k_{\mathrm{B}}T}{\gamma} + C\,\mathrm{e}^{-(\gamma/m)t}, \tag{9.15}$$

where C is a constant that depends on the initial conditions. If one starts from equilibrium, one has $C = -2k_B T/\gamma$, so that $z(t) \simeq 2\langle v^2(0)\rangle\, t$ for small values of t. After some time τ_r of the order of m/γ, the displacement squared $\Delta x^2(t)$ of the Brownian particle grows approximately linearly with t on average:

$$\frac{1}{2}\lim_{t\to\infty}\frac{\mathrm{d}\langle\Delta x^2(t)\rangle}{\mathrm{d}t} = D = \frac{k_B T}{\gamma}. \tag{9.16}$$

The constant D that appears in this relation is called the **diffusion coefficient**. This relation between D and γ is known as the **Einstein relation**. From what has preceded, one can see that the Einstein relation allows one to evaluate $k_B T$ if one follows the displacements of a Brownian particle of known radius R. This is the most surprising result of Einstein's 1905 article. (It is interesting that this relation had been obtained in 1904 by the Australian physicist W. Sutherland [Suth05] in a paper that was eventually published almost at the same time as Einstein's.) Let us note that D is independent of the mass m of the Brownian particle, and depends only on its radius R (via eq. (9.1)), on the viscosity η of the solvent, and on temperature, agreeing with Perrin's qualitative observations. Only the relaxation time $\tau_r = m/\gamma$ depends on m, but it is usually too small to be accessible by observing the particle displacements through a microscope. Indeed, for a particle with a radius of 1 μm, one has $D \simeq 2.6 \cdot 10^{-13}\,\mathrm{m^2 s^{-1}}$, which corresponds to a displacement of about 5.6 μm per minute. The particles used by Perrin in his experiments had a density equal to about 1.2 times that of water, and a radius of the order of magnitude of 1 μm. For such particles, the relaxation time τ_r is of the order of $2.6 \cdot 10^{-10}$ s, and the transients are therefore unobservable.

An equivalent expression of the diffusion coefficient D can be obtained by considering the velocity correlation function. By multiplying both sides of eq. (9.4) by $v(0)$ and taking the average, we obtain

$$m\frac{\mathrm{d}}{\mathrm{d}t}\langle v(t)v(0)\rangle = -\gamma\langle v(t)v(0)\rangle + \langle f(t)v(0)\rangle. \tag{9.17}$$

The last term vanishes for $t > 0$. On the other hand,

$$\langle v^2(0)\rangle = \frac{k_B T}{m}, \tag{9.18}$$

because of the equipartition theorem, which provides us with the initial condition. We thus obtain

$$\langle v(t)v(0)\rangle = \frac{k_B T}{m}\mathrm{e}^{-(\gamma/m)t}, \qquad t \geq 0. \tag{9.19}$$

At equilibrium, the statistical properties of the system should not depend on the absolute time. We can then add to the time abscissa an arbitrary amount t', and averages and correlations should remain unchanged. This property is known as **time-translation invariance**. We thus have, for example,

$$\langle v(t+t')v(t')\rangle = \langle v(t)v(0)\rangle, \tag{9.20}$$

for any value of t'. More particularly, by setting $t' = -t$, we obtain

$$\langle v(t)v(0)\rangle = \langle v(0)v(-t)\rangle, \tag{9.21}$$

which allows us to extend the expression of the correlation function to include negative times. We thus obtain

$$\langle v(t)v(0)\rangle = \frac{k_B T}{m} e^{-(\gamma/m)|t|}, \qquad \forall t. \tag{9.22}$$

Let us now consider the definition of the diffusion coefficient that follows from eq. (9.16):

$$D = \frac{1}{2} \lim_{t\to\infty} \frac{d\langle \Delta x^2(t)\rangle}{dt}. \tag{9.23}$$

Since $\Delta x(t) = \int_0^t dt'\, v(t')$, we obtain

$$\langle \Delta x^2(t)\rangle = \int_0^t dt' \int_0^t dt''\, \langle v(t')v(t'')\rangle. \tag{9.24}$$

By taking the derivative of this expression with respect to t, we obtain

$$\frac{d}{dt}\langle \Delta x^2(t)\rangle = 2\int_0^t dt'\, \langle v(t)v(t')\rangle, \tag{9.25}$$

and therefore

$$D = \frac{1}{2} \lim_{t\to\infty} \frac{d\langle \Delta x^2(t)\rangle}{dt} = \lim_{t\to\infty} \int_0^t dt'\, \langle v(t)v(t')\rangle. \tag{9.26}$$

By exploiting time-translation invariance, one can rewrite this relation in the form

$$D = \int_0^\infty dt\, \langle v(t)v(0)\rangle. \tag{9.27}$$

This expression, which connects a kinetic coefficient (D) to a dynamic correlation function (in our case, of v), is a first example of a class of relations known as **Green-Kubo formulas**, which we discuss in more detail in sec. 9.7.

It is now easy to describe Brownian motion in d dimensions. It is useful to define the diffusion coefficient D as follows:

$$D = \frac{1}{2d} \lim_{t\to\infty} \frac{d\langle \Delta r^2(t)\rangle}{dt} = \frac{1}{2d} \lim_{t\to\infty} \frac{d}{dt} \sum_{\alpha=1}^{d} \langle \Delta r_\alpha^2(t)\rangle, \tag{9.28}$$

where $\Delta \boldsymbol{r}(t) = (\Delta r_\alpha(t))$, $\alpha = 1, \ldots, d$, is the displacement vector of the Brownian particle from its position at time $t = 0$. The Einstein relation between D and γ thus remains unchanged, and the same applies for that between Λ and γ, taking into account eq. (9.3) and the equipartition theorem. The d components of the displacement are mutually independent. Relation (9.27) is modified as follows:

$$D = \frac{1}{d} \int_0^\infty dt\, \langle \boldsymbol{v}(t) \cdot \boldsymbol{v}(0)\rangle. \tag{9.29}$$

Proving these results becomes an easy exercise for the reader.

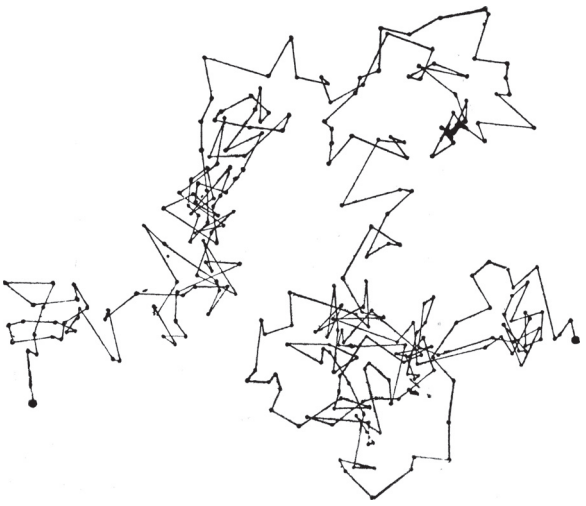


FIGURE 9.1. An experimental trajectory of a Brownian particle. From (Perr16, fig. 8).

9.3 Fractal Properties of Brownian Trajectories

Perrin's experiments, as described in [Perr16], used particles of gamboge, a hard, brittle gum resin that is obtained from various Southeast Asian trees of the genus Garcinia. Particles of approximately uniform size were painstakingly selected via fractioned centrifugation, and their density was accurately measured by three different methods, obtaining agreement to the third decimal figure. Similarly, their volume was evaluated by several methods. The suspension of gamboge particles was observed in a layer about a tenth of a millimeter thick, via a camera lucida, a device that projects an image on a plane surface, allowing the tracking of the particles. We can see in figure 9.1 the extremely irregular trajectory traveled by a Brownian particle observed by Perrin [Perr16, p. 115–116]. He remarks that

> As a matter of fact diagrams of this sort, and even the next figure, in which a large number of displacements are traced on an arbitrary scale, give only a very meager idea of the extraordinary irregularity of the actual trajectory. For if the positions were to be marked at intervals of time 100 times shorter, each segment would be replaced by a polygonal contour relatively just as complicated as the whole figure, and so on. Obviously it becomes meaningless to speak of a tangent to a trajectory of this kind.

As the time interval over which the displacement is recorded becomes smaller, the irregularity of the trajectory increases, as shown in fig. 9.2. Here, we reproduce the position of a simulated particle over 64 successive time intervals equal to τ (the full line). In the same figure, the thin line shows positions measured over 2,048 time intervals, each one lasting $\tau/32$. In this manner, one can observe the progressive increase in complexity of the trajectory. Only if we were able to observe the displacements of the particle over times on the order of τ_r, the correlation time of the velocities, would the Brownian trajectory begin to look smooth. As long as the interval τ between successive observations remains much

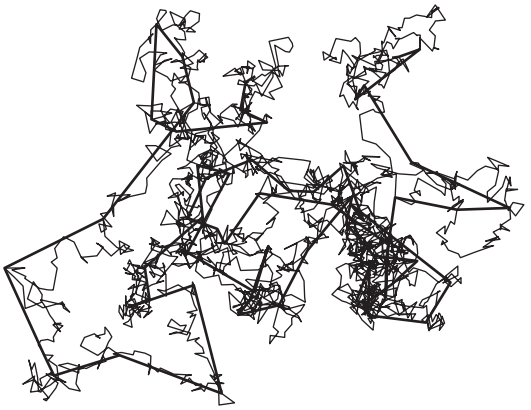


FIGURE 9.2. A simulated Brownian trajectory described with increasing resolution. The full line represents the trajectory with 64 positions taken at time intervals τ. The thin line represents 2,048 positions at time intervals equal to $\tau/32$.

larger than τ_r, the velocity estimate obtained from the $\Delta x/\tau$ ratio will tend to grow as $\tau^{-1/2}$ as τ decreases. This paradoxical behavior had also been noticed by Gouy [Gouy88] and it represented one of the obstacles to the understanding of Brownian motion that Einstein so brilliantly overcame. Of course, this holds only as long as τ is much larger than the relaxation time τ_r.

The Brownian trajectory (at least as long as it is not measured over time intervals on the order of τ_r) is therefore a **fractal**, because it possesses scale-invariance properties: one of its segments, suitably enlarged, has the same statistical properties as the entire trajectory. L. F. Richardson pointed out in 1961 that the measured length of coastline of the land frontier of several countries [Rich61] increased in a power-law way when the resolution used to describe it decreased. Systematizing this observation [Mand67], B. Mandelbrot found that this behavior could be described in terms of fractional dimensions, and introduced the concept of fractal [Mand77] to designate objects with extremely irregular shapes, which, however, possess the property of scale invariance. Historically, the first example of a fractal object (well before the introduction of the term) was precisely Brownian trajectories.

Let us suppose that we want to measure the length of the path traveled by a Brownian particle during a certain time interval T. We subdivide this interval into subintervals of duration $\tau = T/n$, and we measure the particle's displacement $\Delta \boldsymbol{r}_i$ in each of these intervals. The total length L_n of the path measured with this degree of resolution is given by

$$L_n = \sum_{i=1}^{n} \sqrt{\Delta r_i^2}. \tag{9.30}$$

The length L_n of the path is obviously a random quantity but, since we know that $\langle \Delta r_i^2 \rangle = 2dD\tau$, we expect that the typical value of $\sqrt{\Delta r_i^2}$ is of order $\sqrt{2dD\tau} = \sqrt{2dDT}/n^{1/2}$. Therefore

$$L_n \propto n \cdot n^{-1/2} = n^{1/2}. \tag{9.31}$$

We see that as the resolution with which we describe the trajectory increases, the run's length increases! This differs from what happens when we are dealing with smooth curves—in this case, the total length tends to a limit, for $n \to \infty$, that is the length of the curve.

The manner in which length grows as resolution increases is described by the **fractal dimension** of the trajectory. An intuitive notion of the dimension of an object is obtained by counting how the number of objects of size ℓ necessary to cover the object in question grows with ℓ. If we consider a smooth and finite curve in a plane, the number n_ℓ of circles of radius ℓ necessary to cover it will vary as L/ℓ and grows therefore as ℓ^{-D_c}, where $D_c = 1$. On the other hand, if we want to cover a regular plane figure (a polygon, for example) with circles of radius ℓ, the number of circles necessary will vary as ℓ^{-D_c}, with $D_c = 2$. We can therefore identify D_c with the dimension $D = 1$ for the line and $D = 2$ for the plane surface. We call D_c the **coverage dimension** of the figure in question.

Let us now assume that we have an infinitely detailed picture of the Brownian trajectory, and ask ourselves how many circles of radius ℓ are necessary to cover it. Since the typical displacement over a time interval τ is proportional to $\sqrt{\tau}$, we can associate a time interval $\tau \propto \ell^2$ with each ℓ, during which the displacement will be on the order of ℓ. In that case, it is clear that the number of circles necessary to cover the trajectory varies as τ^{-1}, and therefore as ℓ^{-D_c} with $D_c = 2$. We thus arrive at the result that *the coverage dimension of the Brownian trajectory is equal to 2*. This result is independent of the dimension of the space in which the Brownian motion occurs, as long as it is at least equal to 2. The Brownian trajectory is therefore a line with fractal dimension equal to 2.[2]

In the same line of reasoning, we can attempt to evaluate the speed of the Brownian particle by measuring its displacement $\Delta r(\tau)$ over a time interval of duration τ. We obtain the estimate $v(\tau) = \Delta r(\tau)/\tau \propto \tau^{-1/2}$. This holds of course only as long as $\tau \gg \tau_r$, where $\tau_r = m/\gamma$ is the relaxation time. But since $\tau_r \sim 10^{-10}$ s for the Brownian particles we are considering, the velocity $v(\tau)$ measured over time intervals accessible to direct observation corresponds to no objective property of the motion under investigation, as emphasized by Einstein [Eins07a].

9.4 Smoluchowski Equation

Let us now consider the probability density $P(\boldsymbol{r}, t)$ that the Brownian particle is located close to the point $\boldsymbol{r}$ at instant t. In order to simplify our reasoning, let us once again limit ourselves to one dimension, and let us denote the only coordinate of $\boldsymbol{r}$ with x.

As a consequence of the law of total probabilities, we can evaluate the probability density $P(x, t + \Delta t)$ that the particle is found close to x at time $t + \Delta t$ by summing over all its possible positions x' at time t. Then $P(x, t + \Delta t)$ is given by the integral over x' of the probability density $P(x', t)$ that the particle is close to x' at time t, times the

[2] There are various ways to generalize the concept of dimension of fractal objects, and *coverage dimension* is only one of them. These ways are not equivalent, and for this reason, it is best to use the concept of *fractal dimension* in an informal sense, and each time specify which particular concept of dimension we are referring to.

conditional probability $P(x, t+\Delta t|x', t)$ that, starting from x' at time t, the particle reaches x at time $t+\Delta t$:

$$P(x, t+\Delta t) = \int \mathrm{d}x'\, P(x, t+\Delta t|x', t)\, P(x', t). \tag{9.32}$$

In a time interval of duration $\Delta t \gg \tau_r$, the Brownian particle undergoes a displacement $\Delta x = x - x'$ whose average vanishes and whose variance is equal to $2D\,\Delta t$. Let us denote by $\phi(\Delta x)$ the distribution of this displacement:

$$\phi(\Delta x) = P(x+\Delta x, t+\Delta t|x, t), \qquad \forall x, t. \tag{9.33}$$

If Δt is small enough (but still much larger than τ_r), the typical values of Δx will be small, and therefore we can expand $P(x', t)$ on the right-hand side of equation (9.32) in a Taylor series in powers of Δx around x. Changing the integration variable to Δx, we obtain

$$P(x, t+\Delta t) = \int \mathrm{d}\Delta x\, \phi(\Delta x) \left[P(x,t) - \Delta x \left.\frac{\partial P}{\partial x'}\right|_{x'=x} + \frac{1}{2}\Delta x^2 \left.\frac{\partial^2 P}{\partial x'^2}\right|_{x'=x} + \cdots \right]. \tag{9.34}$$

The integral of the second term in brackets vanishes because $\langle \Delta x \rangle = 0$. To evaluate the integral of the third term, we exploit the relation

$$\langle \Delta x^2 \rangle = \int \mathrm{d}\Delta x\, \phi(\Delta x)\, \Delta x^2 = 2D\,\Delta t. \tag{9.35}$$

We thus obtain

$$P(x, t+\Delta t) = P(x,t) + D\,\Delta t \frac{\partial^2 P}{\partial x^2} + \mathrm{o}(\Delta t). \tag{9.36}$$

Going to the $\Delta t \to 0$ limit, we see that $P(x, t)$ satisfies the **diffusion equation** (also called **Fick's equation**):

$$\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial x^2}. \tag{9.37}$$

It is interesting to consider the specific solution of this equation that satisfies the initial condition $P(x, t=0) = \delta(x)$, where $\delta(x)$ is a Dirac delta. It is obviously none other than the distribution of the displacement Δx of the particle during a time interval of duration t. In order to arrive at this solution, let us transform both sides of eq. (9.37) according to Fourier (see Appendix D) to obtain

$$\frac{\partial P_k}{\partial t} = -Dk^2\, P_k. \tag{9.38}$$

Since $P_k(t{=}0) = \int \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}kx}\, \delta(x) = 1$, we obtain

$$P_k(t) = \mathrm{e}^{-Dk^2 t}, \qquad t \geq 0. \tag{9.39}$$

By taking the inverse Fourier transform, using

$$P(x, t) = \int \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{\mathrm{i}kx}\, P_k(t), \tag{9.40}$$

we finally arrive at

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right). \tag{9.41}$$

The distribution of the displacements is therefore a Gaussian. This was to be expected on the basis of the central limit theorem (see sec. A.6), since the displacement over the time interval of duration t is the sum of a large number of small independent displacements.

Let us now suppose that we have a suspension of a very large number (N) of identical Brownian particles, and denote the local density of Brownian articles by $\rho(x,t)$. If the suspension is sufficiently diluted, so that the Brownian particles do not interact among themselves and can thus be considered independent, $\rho(x,t)$ will obey the same equation (9.37). Therefore, if at a certain instant the density $\rho(x,t)$ is not uniform, a flow of particles proportional to its derivative will be produced. This flow is called the **diffusion current** and its density is expressed by

$$j_D = -D\frac{\partial \rho}{\partial x}. \tag{9.42}$$

Einstein [Eins05] remarked that the existence of this current is a necessary consequence of van 't Hoff's theory of the thermodynamics of diluted solutions (cf. [Call85, sec. 13.5]), since there is no reason, from the point of view of the molecular theory of heat, to consider a suspension of Brownian particles as substantially different from a diluted solution.

We can now directly obtain the Einstein relation between the damping coefficient γ and the diffusion constant D, by supposing that the particles are subject, beyond the fluctuating force we considered so far, to a uniform, externally applied force F—for example, in the direction of the positive x axis. As a consequence of this force, the particles will be driven, on average, by a velocity v_{lim} equal to F/γ. There will therefore by a systematic current j_F given by

$$j_F = \rho\, v_{\text{lim}} = \rho\frac{F}{\gamma}. \tag{9.43}$$

At equilibrium, we should have $j_D + j_F = 0$, and therefore ρ_{eq} must satisfy the relation

$$\rho_{\text{eq}}\frac{F}{\gamma} = D\frac{\partial \rho_{\text{eq}}}{\partial x}. \tag{9.44}$$

This equation admits the solution

$$\rho_{\text{eq}}(x) \propto \exp\left(\frac{Fx}{\gamma D}\right). \tag{9.45}$$

On the other hand, $\rho_{\text{eq}}(x)$ is given by the Maxwell-Boltzmann distribution, expressed as a function of the potential $U(x)$ associated with the force F by

$$\rho_{\text{eq}}(x) \propto \left[-\frac{U(x)}{k_{\text{B}}T}\right]. \tag{9.46}$$

Since, for a uniform force F, one has $U(x) = -Fx + \text{const.}$, by comparing eqs. (9.45) and (9.46), we obtain the Einstein relation in the form

$$\gamma D = k_B T, \tag{9.47}$$

where the force F has disappeared.

Exercise 9.1 Defining

$$\langle \Delta r^2(t) \rangle = \int \mathrm{d}\boldsymbol{r}\, r^2\, P(\boldsymbol{r}, t \,|\, 0, 0),$$

show that

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \Delta r^2(t) \rangle = 2dD,$$

where $P(\boldsymbol{r}, t \,|\, 0, 0)$ satisfies the equations

$$\int \mathrm{d}\boldsymbol{r}\, P(\boldsymbol{r}, t \,|\, 0, 0) = 1; \qquad \frac{\partial P}{\partial t} = D\nabla^2 P.$$

Let us now obtain the equation satisfied by $\rho(x, t)$ (or by the single-particle distribution $P(x, t)$) in the presence of a nonuniform force-field $F(x)$, which derives from a potential $U(x)$, and for which therefore $F(x) = -\mathrm{d}U(x)/\mathrm{d}x$. In order to accomplish this, let us generalize the derivation of eq. (9.37), starting from eq. (9.32), which we rewrite as follows:

$$P(x, t + \Delta t) = \int \mathrm{d}x'\, \phi(\Delta x, x')\, P(x', t). \tag{9.48}$$

In the present case, the distribution $\phi(\Delta x, x')$ of the displacement $\Delta x = x - x'$ also depends on the starting point x', since the system is not uniform. To proceed, we introduce an arbitrary function $f(x)$ and evaluate its average with respect to the probability distribution $P(x, t + \Delta t)$:

$$\langle f \rangle_{t+\Delta t} = \int \mathrm{d}x f(x)\, P(x, t + \Delta t) = \int \mathrm{d}x \int \mathrm{d}x' f(x)\, \phi(\Delta x, x')\, P(x', t). \tag{9.49}$$

For each value of x', we expand $f(x)$ in a series of Δx around x', and we change the integration variable to $\Delta x = x' - x$. We thus obtain

$$\begin{aligned} \langle f \rangle_{t+\Delta t} = \int \mathrm{d}\Delta x \int \mathrm{d}x' \left[f(x') + \Delta x f'(x') + \frac{1}{2}\Delta x^2 f''(x') + \cdots \right] \\ \times \phi(\Delta x, x')\, P(x', t). \end{aligned} \tag{9.50}$$

At this point, we can evaluate the integral over Δx, by exploiting the following relations:

$$\langle \Delta x \rangle_{x'} = \int \mathrm{d}\Delta x\, \Delta x\, \phi(\Delta x, x') = \frac{F(x')}{\gamma} \Delta t, \tag{9.51}$$

$$\langle \Delta x^2 \rangle_{x'} = \int \mathrm{d}\Delta x\, \Delta x^2\, \phi(\Delta x, x') = 2D\, \Delta t. \tag{9.52}$$

These relations follow immediately from the preceding considerations. Changing the integration variable, we now have

$$\langle f \rangle_{t+\Delta t} = \int \mathrm{d}x \left[f(x) + \frac{F(x)}{\gamma} \Delta t f'(x) + D\,\Delta t f''(x) + \cdots \right] P(x,t). \tag{9.53}$$

Assuming that $P(x,t) \to 0$ for $x \to \pm\infty$, we can integrate by parts, and obtain

$$\begin{aligned} \langle f \rangle_{t+\Delta t} = \int \mathrm{d}x f(x) \Bigg\{ & P(x,t) + \Delta t \frac{\partial}{\partial x} \left[-\frac{F(x)}{\gamma} P(x,t) \right] \\ & + \Delta t \frac{\partial^2}{\partial x^2} [D\,P(x,t)] + \mathrm{o}(\Delta t) \Bigg\}. \end{aligned} \tag{9.54}$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle f \rangle_t = \int \mathrm{d}x f(x) \left\{ \frac{\partial}{\partial x} \left[-\frac{F(x)}{\gamma} P(x,t) \right] + \frac{\partial^2}{\partial x^2} [D\,P(x,t)] \right\}. \tag{9.55}$$

Since one obviously has

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle f \rangle_t = \int \mathrm{d}x f(x) \frac{\partial}{\partial t} P(x,t), \tag{9.56}$$

and $f(x)$ is arbitrary, we obtain the equation satisfied by $P(x,t)$:

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x} \left[-\frac{F(x)}{\gamma} P + \frac{\partial}{\partial x} (D\,P) \right]. \tag{9.57}$$

This is the **Smoluchowski equation**. It is easy to see that, if γ and D are connected by the Einstein relation (9.47), and $F(x) = -\mathrm{d}U(x)/\mathrm{d}x$, the equation is satisfied identically by the time-independent Maxwell-Boltzmann distribution $\rho_{\mathrm{eq}}(x) \propto \mathrm{e}^{-U(x)/k_{\mathrm{B}}T}$.

By the same reasoning, one obtains that the corresponding equation for Brownian motion in d dimensions—subject to the force $\boldsymbol{F}$, and under the hypothesis that the fluctuating force $f(t)$ satisfies equations (9.2) and (9.3), where Λ is expressed by eq. (9.10)—has the following expression:

$$\frac{\partial P}{\partial t} = \frac{1}{\gamma} \nabla \cdot [-\boldsymbol{F}(\boldsymbol{r})\, P + k_{\mathrm{B}} T \nabla P]. \tag{9.58}$$

In a colloidal suspension of spherical particles of radius R and mass density ρ_m, the force acting on the particles is given by $\boldsymbol{F} = \mathcal{V}(\rho_m - \rho_{\mathrm{H_2O}})\boldsymbol{g}$, where $\mathcal{V} = 4\pi R^3/3$ is the particle volume, $\boldsymbol{g}$ is the gravity acceleration, and $\rho_{\mathrm{H_2O}}$ is the mass density of water. Therefore the particle distribution in the suspension decreases exponentially with the height z:

$$\rho(z) \propto \mathrm{e}^{-z/z_0}, \tag{9.59}$$

where

$$z_0 = \frac{k_{\mathrm{B}} T}{\mathcal{V} g (\rho_m - \rho_{\mathrm{H_2O}})}. \tag{9.60}$$

Since the quantities that appear in this expression, apart from the Boltzmann constant, are all measurable, a careful measurement of the dependence of the particle density on height allows one to determine k_B and therefore the size of molecules. This is the path followed by J. Perrin in his important work on colloidal suspensions, carefully described in [Perr10]. By this approach it is not necessary to measure the viscosity of the solvent, but it is necessary to measure accurately the particle density. At room temperature, the order of magnitude of z_0 is a few tens of μm for particles of radius $\simeq 0.2$ μm and density larger than that of water by about 20%.

9.5 Diffusion Processes and the Fokker-Planck Equation

The considerations reported in the preceding section can be easily extended to more general systems. Let us assume that we are dealing with a physical system whose behavior is described by a collection of random functions of time $X(t)$, where $X = (X_1, \dots, X_n)$, in which the following hypotheses are satisfied:

1. The process is **Markovian**—in other words, the conditional probability

   $$P(x, t|x^{(k)}, t_k; x^{(k-1)}, t_{k-1}; \dots; x^{(0)}, t_0)$$

   that $X(t) = x$, given that $X(t_k) = x^{(k)}, \dots, X(t_0) = x^{(0)}$, with $t_0 < \cdots < t_k$, depends only on the last state $x^{(k)}$ at t_k, and not on the previous history:

   $$P(x, t|x^{(k)}, t_k; x^{(k-1)}, t_{k-1}; \dots; x^{(0)}, t_0) = P(x, t|x^{(k)}, t_k). \tag{9.61}$$

2. The increment ΔX of the process during a time interval $[t, t + \Delta t]$ satisfies the following conditions:

   - The expected value of $\Delta X = X(t + \Delta t) - X(t)$ is proportional to Δt:

     $$\langle \Delta X_\alpha \rangle = b_\alpha(x)\, \Delta t + o(\Delta t), \tag{9.62}$$

     where $x = (X_\alpha(t))$ and where $b_\alpha(x)$, for each $\alpha = 1, \dots, n$, is a function of the n variables $x_1, \dots, x_n$, and possibly of time t.

   - The expected values of the products $\Delta X_\alpha\, \Delta X_\beta$ are also proportional to Δt:

     $$\langle \Delta X_\alpha\, \Delta X_\beta \rangle = 2\Lambda_{\alpha\beta}(x)\, \Delta t + o(\Delta t), \tag{9.63}$$

     where $\Lambda_{\alpha\beta}(x)$ is a symmetric positive semidefinite matrix that depends on x and possibly on time t.

If these conditions are satisfied, the random function $X(t)$ is called a **diffusion process**. By applying the method that we used to derive the Smoluchowski equation, we can see that its probability density $P(x, t)$ satisfies the following equation:

$$\frac{\partial P}{\partial t} = \sum_{\alpha=1}^{n} \frac{\partial}{\partial x_\alpha} \left\{ -b_\alpha(x)\, P(x, t) + \sum_\beta \frac{\partial}{\partial x_\beta} \left[\Lambda_{\alpha\beta}(x)\, P(x, t) \right] \right\}. \tag{9.64}$$

This equation is known as the **Fokker-Planck equation**, or also as the **forward Kolmogorov equation**.

The Fokker-Planck equation can be applied, for example, to describe the motion of a Brownian particle in one dimension, when the inertia of the particle cannot be neglected. In this case, the state of the particle is described by both its position x and its momentum p. We then obtain

$$\begin{aligned}
\langle \Delta x \rangle &= \frac{p}{m}\,\Delta t + \mathrm{o}(\Delta t),\\
\langle \Delta p \rangle &= \left[F(x) - \gamma\,\frac{p}{m}\right]\Delta t + \mathrm{o}(\Delta t),\\
\langle \Delta v^2 \rangle &= 2\Lambda\;\Delta t + \mathrm{o}(\Delta t),\\
\langle \Delta x^2 \rangle &= \langle \Delta x\,\Delta p \rangle = \mathrm{o}(\Delta t).
\end{aligned} \tag{9.65}$$

We thus obtain a diffusion process whose probability density $P(x, p, t)$ satisfies the Fokker-Planck equation

$$\frac{\partial P}{\partial t} = \frac{\partial}{\partial x}\left(-\frac{p}{m}\,P\right) + \frac{\partial}{\partial p}\left\{\left[-F(x) + \gamma\,\frac{p}{m}\right]P\right\} + \gamma\,k_{\mathrm{B}}T\,\frac{\partial^2 P}{\partial p^2}, \tag{9.66}$$

where we have exploited the Einstein relation $\Lambda = 2\gamma\,k_{\mathrm{B}}T$. This equation, which is easy to generalize to d dimensions, is known as the **Klein-Kramers equation**. It is usually not very useful for actual Brownian particles, but it may be used to describe, for instance, the behavior of the mirror in the Kappler experiment, discussed in sec. 4.1 and in the exercise that follows.

Exercise 9.2 (Kappler's experiment) In Kappler's experiment (see sec. 4.1) a mirror with moment of inertia equal to I is suspended from a taut wire. The torque exerted on the mirror for an angular displacement φ from the equilibrium point is given by $-\kappa\,\varphi$. The mirror is enclosed in a room at the temperature T and the resistance torque exerted on it when its angular velocity equals ω is given by $\tau = -\mu\omega$.

1. Write down the Kramers equation governing the evolution of the angular displacement φ from its equilibrium position.
2. By evaluating the Kramers equation for the equilibrium (Boltzmann) distribution, obtain the Einstein relation connecting the noise intensity and the resistance coefficient μ.
3. Write down the Langevin equation satisfied by the angular displacement $\varphi(t)$. Derive from it the differential equation satisfied by the correlation function $C(t) = \langle \varphi(t)\varphi(0) \rangle$ for $t > 0$.
4. Assuming $C(t) = \alpha\,\mathrm{e}^{-\lambda_+ t} + \beta\,\mathrm{e}^{-\lambda_- t}$, where $\lambda_\pm$ are constants to be determined, and suitable initial conditions, give the expression of $C(t)$. Notice that there are two regimes, one for small values of μ, and one for large values of μ. Identify the value μ_{c} of μ separating these regimes.

9.6 Correlation Functions

Let us observe that the preceding results can also be read differently. Let us assume that we are not dealing with a single Brownian particle, but with a distribution of noninteracting Brownian particles, where the local density close to point $\boldsymbol{r}$ is given by $\rho(\boldsymbol{r}, t)$. Equation (9.64) can be rewritten as follows:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \boldsymbol{J}, \tag{9.67}$$

where the particle current $\boldsymbol{J}$ is given by

$$\boldsymbol{J} = -\frac{1}{\gamma}\nabla U - D\nabla \rho. \tag{9.68}$$

The proportionality coefficient between the particle density ρ and the probability density P is obviously the total number of particles N.

Let us return to considering the problem in the absence of an external potential $U(\boldsymbol{r})$ and write the diffusion equation in Fourier space. Defining

$$\delta\rho(\boldsymbol{r}, t) = \rho(\boldsymbol{r}, t) - \langle \rho \rangle, \tag{9.69}$$

where $\langle \rho \rangle$ is the average particle density (assumed uniform), and

$$\rho(\boldsymbol{k}, t) = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\, \delta\rho(\boldsymbol{r}, t), \tag{9.70}$$

we obtain

$$\frac{\partial}{\partial t}\rho(\boldsymbol{k}, t) = -Dk^2\, \rho(\boldsymbol{k}, t), \tag{9.71}$$

where $k = |\boldsymbol{k}|$. A possible perturbation with respect to the uniform equilibrium state therefore disappears in a characteristic time $\Omega(\boldsymbol{k})^{-1}$, where

$$\Omega(\boldsymbol{k}) = Dk^2. \tag{9.72}$$

Obviously, the local particle density $\rho(\boldsymbol{r}, t)$ can also be regarded as a random quantity—since each Brownian particle is subjected to an independent random force, the corresponding density will fluctuate over time. In this case, the Fokker-Planck equation will hold *on average*, for the same reason that the frequency of a random result is, on average, equal to its probability. A moment of reflection, however, shows us that while the continuity equation (9.67) must be valid for the fluctuating density ρ instant by instant, so as to maintain the total number of particles constant, the relation (9.68) between the current $\boldsymbol{J}$ and the density ρ is valid only on average. In fact, this relation is a statistical constraint between $\boldsymbol{J}$ and ρ, which we obtained by taking the average of the displacements $\Delta\boldsymbol{r}$ for all particles found in the neighborhood of $\boldsymbol{r}$. It is therefore a **constitutive relation** between density and current, containing the damping coefficient γ, itself related to the viscosity of the solvent.

We can thus consider $\delta\rho(\boldsymbol{r}, t)$ as a random function of $\boldsymbol{r}$ and t, whose average vanishes. Let us now evaluate the correlation function

$$C(\boldsymbol{r}, t) = \langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(0, 0) \rangle . \tag{9.73}$$

It satisfies the equation

$$\frac{\partial}{\partial t} C(\boldsymbol{r}, t) = D\nabla^2 C(\boldsymbol{r}, t). \tag{9.74}$$

Taking the Fourier transform

$$C(\boldsymbol{k}, t) = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\, C(\boldsymbol{r}, t), \tag{9.75}$$

we obtain the equation

$$\frac{\partial}{\partial t} C(\boldsymbol{k}, t) = -Dk^2\, C(\boldsymbol{k}, t), \tag{9.76}$$

for $t \geq 0$. We obtain therefore

$$C(\boldsymbol{k}, t) = \exp\left(-Dk^2 t\right)\, C(\boldsymbol{k}, 0), \qquad t \geq 0. \tag{9.77}$$

Now, we have

$$\begin{aligned} \langle \rho(\boldsymbol{k}, t)\, \rho(-\boldsymbol{k}, 0) \rangle &= \int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(\boldsymbol{r}', 0) \rangle \\ &= \int \mathrm{d}\Delta\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\Delta\boldsymbol{r}} \langle \delta\rho(\boldsymbol{r}' + \Delta\boldsymbol{r}, t)\, \delta\rho(\boldsymbol{r}', 0) \rangle \\ &= \int \mathrm{d}\boldsymbol{r}' \int \mathrm{d}\Delta\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\Delta\boldsymbol{r}}\, C(\Delta\boldsymbol{r}, t) \\ &= V\, C(\boldsymbol{k}, t), \end{aligned} \tag{9.78}$$

where we have exploited translation invariance and V is the volume of the system. Let us assume that the system is at equilibrium and therefore that $\langle \rho(\boldsymbol{k}, t)\rho(-\boldsymbol{k}, 0) \rangle$ satisfies time-translation invariance. We thus obtain

$$\langle \rho(\boldsymbol{k}, -t)\rho(-\boldsymbol{k}, 0) \rangle = \langle \rho(\boldsymbol{k}, 0)\rho(-\boldsymbol{k}, t) \rangle . \tag{9.79}$$

We have therefore, taking into account eq. (9.77),

$$C(\boldsymbol{k}, t) = \exp\left(-Dk^2\, |t|\right)\, C(\boldsymbol{k}, 0). \tag{9.80}$$

We can now take the Fourier transform with respect to t. In the time domain, it is customary to define the transformation by

$$f(\omega) = \int \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t} f(t), \tag{9.81}$$

rather than by the convention used in the space domain:

$$f(k) = \int \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}\omega x} f(x). \tag{9.82}$$

We thus express $\mathrm{c}(\boldsymbol{k}, \omega)$ by

$$C(\boldsymbol{k}, \omega) = \int \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t}\, C(\boldsymbol{k}, t). \tag{9.83}$$

We thus obtain

$$\begin{aligned} C(\boldsymbol{k}, \omega) &= C(\boldsymbol{k}, t{=}0) \left[\int_{-\infty}^{0} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t + Dk^2 t} + \int_{0}^{+\infty} \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t - Dk^2 t} \right] \\ &= C(\boldsymbol{k}, t{=}0) \frac{2Dk^2}{\omega^2 + (Dk^2)^2}. \end{aligned} \tag{9.84}$$

The correlation function $C(\boldsymbol{k}, \omega)$ is therefore a **Lorentzian**, centered at $\omega = 0$, and with a width equal to $\Omega(\boldsymbol{k}) = Dk^2$.

The coefficient $C(k, t{=}0)$ can be related to the equilibrium properties of the system. For each value of t, because of time-translation invariance, one has in fact

$$C(\boldsymbol{k}, t = 0) = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \left\langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(\boldsymbol{r}', t) \right\rangle, \tag{9.85}$$

where $\boldsymbol{r}'$ is arbitrary. Our system of Brownian particles behaves like a diluted gas with average density equal to $\langle\rho\rangle$. At equilibrium, since under our hypothesis the Brownian particles do not interact, we can assume that the density fluctuations at different points in space are independent. More specifically, let us consider an element of volume $\mathrm{d}V$ such that, on average, it contains a very small number of particles $\langle\nu\rangle = \langle\rho\rangle\ \mathrm{d}V$. The number of particles actually contained in this volume obeys a Poisson distribution, and one therefore has $\left\langle (\nu - \langle\nu\rangle)^2 \right\rangle = \langle\nu\rangle$. From this, it follows that the variance in the number of particles N contained in a small but macroscopic volume V is equal to its average N. We conclude that $\rho(\boldsymbol{r}, t)$ is a random Gaussian function of $\boldsymbol{r}$ of average $\langle\rho\rangle$, whose spatial and temporal correlations decay very rapidly. The fact that $\rho(\boldsymbol{r}, t)$ is approximately Gaussian follows once again from the central limit theorem. We can therefore set

$$\left\langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(\boldsymbol{r}', t) \right\rangle \propto \delta(\boldsymbol{r} - \boldsymbol{r}'). \tag{9.86}$$

On the other hand,

$$\int \mathrm{d}\boldsymbol{r} \int \mathrm{d}\boldsymbol{r}'\, \left\langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(\boldsymbol{r}', t) \right\rangle = \left\langle \Delta N^2 \right\rangle. \tag{9.87}$$

However, $\langle \Delta N^2 \rangle = N$, and because of translational invariance, one of the two integrals gives us the sample volume V. Therefore,

$$\left\langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(\boldsymbol{r}', t) \right\rangle = \frac{N}{V} \delta(\boldsymbol{r} - \boldsymbol{r}') = \langle\rho\rangle\, \delta(\boldsymbol{r} - \boldsymbol{r}'). \tag{9.88}$$

By transforming this relation according to Fourier, we obtain

$$C(\boldsymbol{k}, t{=}0) = \int \mathrm{d}\boldsymbol{r}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \left\langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(0, t) \right\rangle = \langle \rho \rangle \,. \tag{9.89}$$

The final expression is therefore

$$C(\boldsymbol{k}, \omega) = \frac{2\,\langle \rho \rangle\; Dk^2}{\omega^2 + (Dk^2)^2}. \tag{9.90}$$

Let us observe that for $\boldsymbol{k} \neq 0$ one has $\lim_{\omega\to 0} C(\boldsymbol{k}, \omega) = 2\,\langle \rho \rangle / Dk^2$, while for $\omega \neq 0$ one has $\lim_{k\to 0} C(\boldsymbol{k}, \omega) = 0$. Therefore $\lim_{k\to 0} \lim_{\omega\to 0} C(\boldsymbol{k}, \omega) = \infty$, whereas $\lim_{\omega\to 0} \lim_{k\to 0} C(\boldsymbol{k}, \omega) = 0$. This noncommutativity of limits appears quite generally in dynamic correlation functions.

9.7 Kubo Formula and Sum Rules

One can relate the diffusion coefficient D to the correlation function of the current. By applying the continuity equation

$$\frac{\partial}{\partial t}\rho + \nabla \cdot \boldsymbol{J} = 0 \tag{9.91}$$

to the correlation function $\left\langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(\boldsymbol{r}', t') \right\rangle$, one obtains

$$\frac{\partial}{\partial t}\frac{\partial}{\partial t'} \left\langle \delta\rho(\boldsymbol{r}, t)\, \delta\rho(\boldsymbol{r}', t') \right\rangle = \sum_{\alpha\beta} \frac{\partial}{\partial \boldsymbol{r}_\alpha}\frac{\partial}{\partial \boldsymbol{r}_\beta} \left\langle J_\alpha(\boldsymbol{r}, t)\, J_\beta(\boldsymbol{r}', t') \right\rangle. \tag{9.92}$$

By transforming according to Fourier, and taking space- and time-translation invariance into account, we obtain

$$\omega^2 \left\langle \rho(\boldsymbol{k}, \omega)\, \rho(-\boldsymbol{k}, -\omega) \right\rangle = \sum_{\alpha\beta} k_\alpha k_\beta \left\langle J_\alpha(\boldsymbol{k}, \omega)\, J_\beta(-\boldsymbol{k}, -\omega) \right\rangle. \tag{9.93}$$

Due to isotropy, the correlation function $\left\langle J_\alpha(\boldsymbol{k}, \omega)\, J_\beta(-\boldsymbol{k}, -\omega) \right\rangle$ can be expanded into a scalar term (which is independent of the direction of $\boldsymbol{k}$), a tensorial quantity (which transforms as the collection of products $k_\alpha k_\beta$), and higher-order terms:

$$\begin{aligned} \left\langle J_\alpha(\boldsymbol{k}, \omega)\, J_\beta(-\boldsymbol{k}, -\omega) \right\rangle = {} & \delta_{\alpha\beta} \left(a_0 + a_1 k^2 + \cdots \right) \\ & + k_\alpha k_\beta \left(b_0 + b_1 k^2 + \cdots \right) + \cdots . \end{aligned} \tag{9.94}$$

Therefore, limiting ourselves to second order in k, we obtain

$$\sum_{\alpha\beta} k_\alpha k_\beta \left\langle J_\alpha(\boldsymbol{k}, \omega)\, J_\beta(-\boldsymbol{k}, -\omega) \right\rangle = a_0 k^2 + \mathrm{O}(k^4). \tag{9.95}$$

On the other hand, by taking the trace of both sides of eq. (9.94), one has

$$a_0 = \frac{1}{d} \lim_{k\to 0} \left\langle \boldsymbol{J}(\boldsymbol{k}, \omega) \cdot \boldsymbol{J}(-\boldsymbol{k}, -\omega) \right\rangle, \tag{9.96}$$

where d is the space dimensionality. Taking eq. (9.90) into account, we obtain

$$\begin{aligned} D\langle\rho\rangle &= \frac{1}{2}\lim_{\omega\to 0}\lim_{k\to 0}\frac{\omega^2}{k^2}\left\langle\rho(\boldsymbol{k},\omega)\,\rho(-\boldsymbol{k},-\omega)\right\rangle \\ &= \frac{1}{2d}\lim_{\omega\to 0}\lim_{k\to 0}\left\langle \boldsymbol{J}(\boldsymbol{k},\omega)\cdot\boldsymbol{J}(-\boldsymbol{k},-\omega)\right\rangle. \end{aligned} \tag{9.97}$$

This equation, which relates the diffusion coefficient to a correlation function, is a new example of the **Kubo formula**. The importance of this relation is that it allows us to calculate dynamic quantities (such as the diffusion coefficient) in terms of correlation functions at thermodynamic equilibrium.

Another relation, apparently trivial but extremely important in the applications, comes from the observation that the correlation function $C(\boldsymbol{k},0)$ for equal times is obtained as an inverse Fourier of the dynamic correlation function $C(\boldsymbol{k},\omega)$:

$$C(\boldsymbol{k},0) = \int \frac{\mathrm{d}\omega}{2\pi}\, C(\boldsymbol{k},\omega). \tag{9.98}$$

The reader is invited to check that this relation is actually valid for our eq. (9.84). The interest in this **sum rule** is connected especially to those cases in which the dynamic correlation function is not perfectly known. In these cases, it imposes an important constraint on the approximations.

Exercise 9.3 From the relation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\Delta r^2(t)\right\rangle = 2\int_0^t \mathrm{d}t'\,\left\langle \boldsymbol{v}(t')\cdot\boldsymbol{v}(0)\right\rangle,$$

assuming

$$\left\langle \boldsymbol{v}(t)\cdot\boldsymbol{v}(0)\right\rangle = \langle v^2\rangle \mathrm{e}^{-t/\tau}, \qquad t>0,$$

evaluate $\langle\Delta r^2(t)\rangle$. Given that the diffusion constant in most liquids is of the order of $10^{-5}\,\mathrm{cm^2 s^{-1}}$, estimate the order of magnitude of τ.

9.8 Metastable States and Stochastic Resonance

Quite often, the transformation of a molecule into a different one in chemical reactions can be described by a one-dimensional diffusion process, that is, the analog of a particle moving in one dimension in a potential $U(x)$ and subject to thermal noise. It becomes then interesting to study how the particle's dwell time in a given metastable state, that is, in a local potential minimum higher than the absolute one, depends on the height of the potential barrier to be overcome.

Consider thus a particle in one dimension, whose coordinate x obeys the Langevin equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\Gamma\frac{\mathrm{d}U(x)}{\mathrm{d}t} + f(t),$$

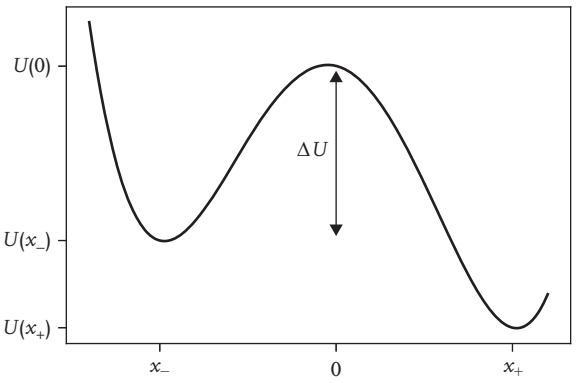


FIGURE 9.3. Potential $U(x)$ for a metastable state x_-. The barrier height $U(0) - U(x_-)$ is denoted by ΔU.

where the Gaussian random force $f(t)$ satisfies the relations

$$\langle f(t)\rangle = 0, \qquad \langle f(t) f(t')\rangle = 2\Gamma k_B T\, \delta(t-t'), \quad \forall t, t'.$$

The potential $U(x)$ is given by

$$U(x) = -\frac{1}{2}\kappa x^2\left(x_0^2 - \frac{x^2}{2}\right) - \kappa A x, \qquad \kappa, A > 0. \tag{9.99}$$

Assuming $A \ll x_0^3$, we see that the potential has two local minima close to $\pm x_0$, and a local maximum close to 0, as shown in fig. 9.3. The value of the potential at the point $x_- \simeq -x_0$ is larger than its value at the point $x_+ \simeq +x_0$. Therefore, if the system is initially placed at x_-, we expect it to cross the barrier at the origin and to reach eventually x_+. We now evaluate, following Kramers [Kram40], the **escape rate** from this **metastable state**, that is, the probability per unit time that the system leaves the metastable region at the left of the origin.

The Fokker-Planck equation for this system reads

$$\frac{\partial p(x,t)}{\partial t} = \Gamma \frac{\partial}{\partial x}\left[U'(x)\, p + k_B T \frac{\partial p}{\partial x}\right]. \tag{9.100}$$

We look for a stationary solution to this equation. Thus the right-hand side of this equation must vanish, and, by integration, we obtain the equation

$$U'(x)\, p + k_B T \frac{\mathrm{d}p}{\mathrm{d}x} = -\frac{J k_B T}{\Gamma} = \text{const.}, \tag{9.101}$$

which defines the probability current J in the steady state, which is equal to the rate by which the particle crosses the barrier at $x = 0$. One can easily check that, in our case, one has

$$J = -\Gamma\, k_B T\, \mathrm{e}^{-U(x)/k_B T} \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{U(x)/k_B T}\, p(x)\right). \tag{9.102}$$

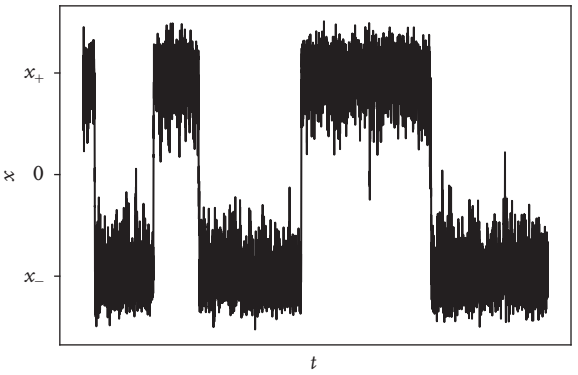


FIGURE 9.4. Behavior of a Brownian particle in a bistable potential in the absence of modulation.

Therefore, given arbitrarily x_1 and x_2, one has

$$J \cdot \int_{x_1}^{x_2} \mathrm{d}x \, \mathrm{e}^{U(x)/k_\mathrm{B}T} = -\Gamma k_\mathrm{B} T \left[\mathrm{e}^{U(x)/k_\mathrm{B}T} \, p(x) \right]_{x_1}^{x_2} . \tag{9.103}$$

Let us take $x_1 = x_-$ and $x_2 = x_+$, and assume that $p(x_+)$ is very small. We then have

$$J \simeq \Gamma \, k_\mathrm{B} T \, \mathrm{e}^{U(x_-)/k_\mathrm{B}T} p(x_-) \Big/ \left(\int_{x_-}^{x_+} \mathrm{d}x \, \mathrm{e}^{U(x)/k_\mathrm{B}T} \right). \tag{9.104}$$

We can obtain $p(x_-)$ by supposing that it is close to an equilibrium distribution for x close to x_-, and by imposing the normalization condition

$$\int_{-\infty}^{+\infty} \mathrm{d}x \, p(x_-) \, \mathrm{e}^{-(U(x)-U(x_-))/k_\mathrm{B}T} \simeq p(x_-) \left(\frac{2\pi k_\mathrm{B} T}{\omega_-^2} \right)^{1/2} = 1, \tag{9.105}$$

where we have supposed $U(x) = U(x_-) + \frac{1}{2}\omega_-^2 (x - x_-)^2$ for x close to x_-, and we have extended the integration limits to infinity, assuming that the probability distribution is concentrated near x_-. The integral in the denominator of eq. (9.104) is dominated by the values of x close to 0, where we have $U(x) \simeq -\frac{1}{2}\omega_0^2 x^2$, and thus, as shown by a straightforward integration, it is equal to $(2\pi k_\mathrm{B} T/\omega_0^2)^{1/2}$. By collecting all these results, we obtain

$$J = \Gamma \, \frac{\omega_0 \, \omega_-}{2\pi} \, \mathrm{e}^{-\Delta U/k_\mathrm{B}T}, \tag{9.106}$$

where $\Delta U = U(0) - U(x_-)$ is the height of the energy barrier. This characteristic behavior as the exponential of minus the barrier height (in units of $k_\mathrm{B} T$) is known as the **Arrhenius law**.

The noise-induced escape over a potential barrier has far-reaching consequences, also in other fields. On the one hand, there is no need to suppose that the point x_+ is more stable than x_-. Suppose, for instance, that $U(x_+) = U(x_-)$. Then, as long as the barrier ΔU is large enough, the system will switch from x_- to x_+ or back, at random times separated, on average, by $1/J$. Such a behavior is shown in fig. 9.4. On the other hand, suppose that the quantity A that appears in eq. (9.99), instead of being constant, is sinusoidally modulated

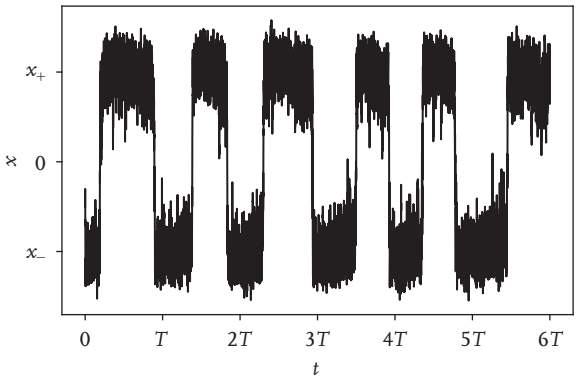


FIGURE 9.5. Behavior of a Brownian particle in a bistable potential modulated with period T, exhibiting stochastic resonance.

with angular frequency Ω, that is, that $A(t) = A_0 \cos \Omega t$. Then, when $t = 0$, the point x_- is metastable and the point x_+ is stable, and the escape rates from x_- and x_+ are respectively given by

$$\begin{aligned} J_- &\simeq \Gamma \frac{\kappa x_0^2}{\sqrt{2}\pi} \exp\left(-\frac{\kappa}{k_B T}\left(\frac{x_0^4}{4} - A_0 x_0\right)\right), \\ J_+ &\simeq \Gamma \frac{\kappa x_0^2}{\sqrt{2}\pi} \exp\left(-\frac{\kappa}{k_B T}\left(\frac{x_0^4}{4} + A_0 x_0\right)\right). \end{aligned} \tag{9.107}$$

When $t = \pi/\Omega$, the two points swap their stability. Suppose now that Ω is such that

$$J_+ < \Omega < J_-, \tag{9.108}$$

where the escape rates are given by the expressions above. Then, if the system is in the temporarily metastable state, say x_-, it has time to overcome the barrier and go to the stable one, x_+. After a half-period, x_+ becomes itself metastable, and the system has time to cross the barrier in the reverse direction.

In other words, assume that we apply to the system a small periodic modulation of amplitude $A_0 \ll x_0^3$ and angular frequency Ω. If the temperature is too small, the system does never have the time to escape from the metastable state during the half-period. If the temperature is too large, the system has the time to make a number of comings and goings from one minimum to the other in each half-period. In neither case does the resulting behavior exhibit oscillations with frequency close to Ω. But, if the temperature satisfies the inequalities

$$\frac{\kappa(x_0^4/4 - A_0 x_0)}{\ln(\sqrt{2}\pi\,\Omega/\,\Gamma\kappa x_0^2)} < k_B T < \frac{\kappa(x_0^4/4 + A_0 x_0)}{\ln(\sqrt{2}\pi\,\Omega/\,\Gamma\kappa x_0^2)}, \tag{9.109}$$

the system will have time to do exactly one crossing, on average, per half-period, and the trajectory will exhibit approximate periodicity with the angular frequency Ω. This behavior is shown in fig. 9.5. Thus, surprisingly, a moderate amount of noise is able to reveal a weak modulation of the potential. This phenomenon, called **stochastic resonance**, applies

of course to more general cases, where the noise does not originate from heat, and its amplitude is not related to temperature. It was in fact discovered in a study of climatic changes in the last 700,000 years. This work [Benz82] was mentioned in the motivation of the Nobel Prize of G. Parisi. It has since found application in several different domains, as discussed, for example, in [Gamm98].

9.9 Chemical Kinetics

The unfolding of chemical reactions is described by the time evolution of the number N_A, $N_B, \ldots$ of molecules of the different chemical species taking part in them. (By **species**, we mean either individual molecules of different kinds, or bound states of more molecules.) For example, consider a reaction where one molecule of type A and one of type B form a bound state AB that can also dissociate into its components. As discussed in sec. 4.12, such a reaction can be represented by the reaction scheme

$$A + B \leftrightharpoons AB. \tag{9.110}$$

Thus the evolution of this system is represented by $(N_A(t), N_B(t), N_{AB}(t))$. The advancement of the reaction from the initial state can be measured by the **extent of reaction** ξ. If the reaction progresses by, say, the formation of a certain number ΔN_{AB} of the complex AB, the extent of reaction increases by

$$\Delta\xi = \Delta N_{AB} = -\Delta N_A = -\Delta N_B. \tag{9.111}$$

In a general case, where the reaction can be written in the form

$$\sum_{i=1}^{r} \nu_i X_i = 0, \tag{9.112}$$

where the stoichiometric coefficients ν_i are negative for the reactants (on the left-hand side of the reaction) and positive for the products (on the right-hand side of the reaction), the change in the extent of reaction is given by

$$\Delta\xi = \frac{\Delta N_i}{\nu_i}. \tag{9.113}$$

This quantity is positive if the reaction proceeds from the reactants to the products, that is, from left to right, and negative otherwise.

The chemical reactions are ultimately due to the collisions of the reactant molecules. If the molecules are diluted enough to be considered independent, the rates of these collisions depend on the molecular concentrations. If these concentrations are uniform (for example, in a well-stirred container) the reaction rates satisfy the **law of mass action**, that is, they are proportional to the product of the reactant concentrations, each raised to the power of the corresponding stochastic coefficient (taken positive). Thus, in the reaction described by eq. (9.110), the rate of AB production is proportional to [A][B] and that of AB dissociation is proportional to [AB], where [X] is the concentration of the chemical species

X. Similarly, in a dimerization reaction of the form $2A \rightleftharpoons AA$, the rate of AA production is proportional to $[A]^2$ and that of AA dissociation is proportional to [AA].

If the system is large enough, so that we can neglect fluctuations in the molecule numbers, the concentrations satisfy a system of ordinary differential equations, called the **rate equations**. For example, for the reaction described in eq. (9.110), the rate equations read

$$\begin{aligned}
\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} &= -k_{\rightarrow}[\mathrm{A}][\mathrm{B}] + k_{\leftarrow}[\mathrm{AB}], \\
\frac{\mathrm{d}[\mathrm{B}]}{\mathrm{d}t} &= -k_{\rightarrow}[\mathrm{A}][\mathrm{B}] + k_{\leftarrow}[\mathrm{AB}], \\
\frac{\mathrm{d}[\mathrm{AB}]}{\mathrm{d}t} &= k_{\rightarrow}[\mathrm{A}][\mathrm{B}] - k_{\leftarrow}[\mathrm{AB}].
\end{aligned} \tag{9.114}$$

Here $k_{\rightarrow}$ and $k_{\leftarrow}$ are the **rate constants** (which depend on temperature). These equations are not independent, since the changes ΔN_i ($i \in \{\mathrm{A}, \mathrm{B}, \mathrm{AB}\}$) of the different molecular species must satisfy condition (4.242), which we write in the form

$$\sum_{i=1}^{k} \nu_i \Delta N_i = \sum_{j=k+1}^{r} \nu_j \Delta N_j, \tag{9.115}$$

where the left-hand side refers to the reactants, and the right-hand side to the products, and where the stoichiometric coefficients are all positive. Therefore the rate equations boil down to a single differential equation describing the evolution of the extent of the reaction ξ, which for our case reads

$$\frac{\mathrm{d}\xi}{\mathrm{d}t} = k_{\rightarrow}[\mathrm{A}][\mathrm{B}] - k_{\rightarrow}[\mathrm{AB}], \tag{9.116}$$

where

$$[\mathrm{A}](t) = [\mathrm{A}](0) - \xi, \quad [\mathrm{B}](t) = [\mathrm{B}](0) - \xi, \quad [\mathrm{AB}](t) = [\mathrm{AB}](0) + \xi. \tag{9.117}$$

Thermodynamics imposes further constraints on the rate constants. Indeed, as shown in sec. 4.12, the total variation $\mathrm{d}G$ of the Gibbs free energy must vanish at equilibrium for an arbitrarily small progress of the reaction. This yields a relation among the chemical potentials of the species involved in the reaction, given by eq. (4.243), which we write in the form

$$-\sum_{i=1}^{k} \nu_i \mu_i^{\mathrm{eq}} + \sum_{j=k+1}^{r} \nu_j \mu_j^{\mathrm{eq}} = 0, \tag{9.118}$$

where μ_i^{eq} is the chemical potential, at equilibrium, of the chemical species X_i. On the other hand, if the law of mass action holds, we must have, at equilibrium, $\mathrm{d}\xi/\mathrm{d}t = 0$, corresponding to

$$k_{\rightarrow} \prod_{i=1}^{k} [\mathrm{X}_i]_{\mathrm{eq}}^{\nu_i} = k_{\leftarrow} \prod_{j=k+1}^{r} [\mathrm{X}_j]_{\mathrm{eq}}^{\nu_j}, \tag{9.119}$$

which implies

$$\frac{k_{\rightarrow}}{k_{\leftarrow}} = \frac{\prod_{j=k+1}^{r} [\mathrm{X}_j]_{\mathrm{eq}}^{\nu_j}}{\prod_{i=1}^{k} [\mathrm{X}_i]_{\mathrm{eq}}^{\nu_i}}. \tag{9.120}$$

Now, for a dilute gas, the chemical potential of a species X_i is given by eq. (4.244), which we write as

$$\mu_i = \mu_i^0 + k_{\mathrm{B}} T \ln[\mathrm{X}_i], \tag{9.121}$$

where μ_i^0 is the "intrinsic" chemical potential (depending, for example, on the interior degrees of freedom). We obtain therefore

$$[\mathrm{X}_i]_{\mathrm{eq}} = \exp\left(\frac{\mu_i^{\mathrm{eq}} - \mu_i^0}{k_{\mathrm{B}} T}\right). \tag{9.122}$$

Substituting in eq. (9.120) and taking into account eq. (9.118) we obtain

$$\frac{k_{\rightarrow}}{k_{\leftarrow}} = \exp\left[\left(\sum_{i=1}^{k} \nu_i \mu_i^0 - \sum_{j=k+1}^{r} \nu_j \mu_j^0\right) \Big/ k_{\mathrm{B}} T\right]. \tag{9.123}$$

Thus the ratio of the forward and backward rate constants depends on the "intrinsic" chemical potentials of the molecules involved in the reaction.

In order to describe chemical reactions, it is often useful to introduce an abstract one-dimensional coordinate that describes the position of a molecular complex along a reaction pathway. This is known as a **reaction coordinate**, or, in a different context, as a **collective variable**. This coordinate can have a spatial interpretation, such as bond length or bond angle, but may have a more complex interpretation. The coordinates interpolate smoothly from the initial configuration (corresponding, for example, to the reactant) to the final one (corresponding to the products). For example, in the dissociation of molecular hydrogen, one could choose as a reaction coordinate the bond length. Then the reaction can be described at the molecular level as the displacement of the system's reaction coordinate, which undergoes a process that can be described as generalized Brownian motion. Typically one assumes that the reaction coordinate x satisfies a Langevin equation of the form of eq. (9.127), where $U(x)$ is the molecule's availability expressed as a function of the reaction coordinate x.

Exercise 9.4 We describe a reaction chain of the form

$$\mathrm{A} \leftrightharpoons \mathrm{A}^* \leftrightharpoons \mathrm{B}.$$

Here A and B are two stable isomers of the same molecule, and A* is an unstable (energy-rich) transition form. The change from A to A* and then to B is described by a reaction coordinate x that moves in a one-dimensional bistable potential $U(x)$, similar to the one shown in fig. 9.3. Thus the phase space for this degree of freedom is divided into three regions, one close to A, one to A*, and the third one to B. The nonequilibrium concentrations of species in the three regions are $[\mathrm{A}](t)$, $[\mathrm{A}^*](t)$, and $[\mathrm{B}](t)$. The rate equations for

this reaction chain read

$$\frac{\mathrm{d}[\mathrm{A}]}{\mathrm{d}t} = k_{\mathrm{AA}^*}[\mathrm{A}^*] - k_{\mathrm{A}^*\mathrm{A}}[\mathrm{A}];$$
$$\frac{\mathrm{d}[\mathrm{A}^*]}{\mathrm{d}t} = -\left(k_{\mathrm{AA}^*} + k_{\mathrm{BA}^*}\right)[\mathrm{A}^*] + k_{\mathrm{A}^*\mathrm{A}}[\mathrm{A}] + k_{\mathrm{A}^*\mathrm{B}}[\mathrm{B}];$$
$$\frac{\mathrm{d}[\mathrm{B}]}{\mathrm{d}t} = k_{\mathrm{BA}^*}[\mathrm{A}^*] - k_{\mathrm{A}^*\mathrm{B}}[\mathrm{B}].$$

These equations imply the conservation law

$$\mathrm{A} + [\mathrm{A}^*] + [\mathrm{B}] = \text{const.}$$

Here k_{ij} is the rate of going from species j to species i. From detailed balance we have

$$\frac{k_{ij}}{k_{ji}} = \frac{[i]_{\mathrm{eq}}}{[j]_{\mathrm{eq}}} = \mathrm{e}^{-\Delta U_{ij}/k_{\mathrm{B}}T},$$

where (i,j) is either $(\mathrm{A}^*, \mathrm{A})$ or $(\mathrm{A}^*, \mathrm{B})$, and $\Delta U_{ij} = U_i - U_j$. Thus, if $\Delta U_{\mathrm{A}^*\mathrm{A}} \gg k_{\mathrm{B}}T$, one has $k_{\mathrm{A}^*\mathrm{A}} \ll k_{\mathrm{AA}^*}$, and analogously for $k_{\mathrm{A}^*\mathrm{B}}$, k_{BA^*}.

Show that, if $\Delta U_{\mathrm{A}^*\mathrm{A}} \gg k_{\mathrm{B}}T$, and analogously for $\Delta U_{\mathrm{A}^*\mathrm{B}}$, the relaxation is dominated by a single relaxation time, τ, and that

$$\tau^{-1} \approx k_{\mathrm{A}^*\mathrm{A}} \text{ or } k_{\mathrm{A}^*\mathrm{B}}.$$

9.10 Generalized Brownian Motion

Starting in the 1950s, with some works by S. Nakajima, R. Zwanzig, and H. Mori (see [Naka58], [Zwan60], [Mori65]), the theory of generalized Brownian motion was developed, with the goal to provide a general description of the evolution of observables in a thermodynamic system. In this section, we will see how the Brownian motion scheme naturally arises when one wants to describe the evolution of a system not in terms of its microscopic variables (whose number is on the order of Avogadro's constant), but in terms of a limited number of variables (which can, however, be fields defined in each point of space, even if slowly varying).

We will additionally derive a number of useful general relations that are the natural outcome of this approach. We will not attempt to formally justify the description in terms of generalized Brownian motion, on the basis of the projection scheme (see [Kubo85, sec. 2.5]). In fact, it is relatively easy to distinguish between a systematic and a random part in the equation of motion for an observable. On the other hand, it is very difficult to justify, in the general case, all the simplifying assumptions that hold for the random force in Brownian motion, like Gaussian distribution, noncorrelation, and the like.

Let us suppose that we have a thermodynamic system, of which X is an observable. We denote by x the value of X. As a consequence of Einstein's theory of fluctuations (sec. 5.17), the probability that X assumes a specific value x is given, at thermodynamic equilibrium

and for a given value T of temperature, by

$$P_{\mathrm{eq}}(x) \propto \exp\left[-\frac{\mathcal{F}(x)}{k_{\mathrm{B}}T}\right], \tag{9.124}$$

where $\mathcal{F}(x)$ is the availability expressed as a function of x:

$$\mathcal{F}(x) = -k_{\mathrm{B}}T \ln \int \mathrm{d}\nu\, \delta[X(\nu) - x]\, \exp\left[-\frac{\mathcal{H}(\nu)}{k_{\mathrm{B}}T}\right]. \tag{9.125}$$

In this expression, ν denotes the microscopic state of the system (for example, for a system of N classical particles, $\nu = ((\boldsymbol{p}_1, \boldsymbol{r}_1), \ldots, (\boldsymbol{p}_N, \boldsymbol{r}_N))$, where $\boldsymbol{p}_i$ and $\boldsymbol{r}_i$ are respectively the momentum and position vector of particle i) and $\mathcal{H}(\nu)$ is the Hamiltonian. The minimum of $\mathcal{F}(x)$ is reached for the equilibrium value x_{eq} of X and is equal to the Helmholtz free energy:

$$\mathcal{F}(x) \geq \mathcal{F}(x_{\mathrm{eq}}) = F(T) = -k_{\mathrm{B}}T \ln \int \mathrm{d}\nu\, \exp\left[-\frac{\mathcal{H}(\nu)}{k_{\mathrm{B}}T}\right]. \tag{9.126}$$

The apparent contradiction between equations (9.125) and (9.126) is resolved by considering that almost all the available phase space corresponds to the value of x_{eq} in the thermodynamic limit.

In a finite system, X is a random quantity that varies according to the system's evolution. Using the analogy with the theory of Brownian motion, we suppose that it is governed by an evolution equation that contains both a systematic and a random part:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = b(x(t)) + f(t), \tag{9.127}$$

where $x(t)$ is the value of X at time t. Let us therefore make the following hypotheses, which can be valid for a suitable choice of X:

- The quantity X varies over time scales that are long when compared to those for which the random force $f(t)$ varies—this allows us to describe $f(t)$ as a random function that is correlated only over very brief time intervals.
- The random force $f(t)$ is a result of a great number of independent interaction processes between X and the other degrees of freedom of the system—this allows us to describe it as a random Gaussian function, with a vanishing average (since its average is taken into account by the systematic term $b(x)$).
- We additionally suppose (for simplicity, but also because we expect to describe *small* fluctuations) that the correlation of $f(t)$ does not depend on the instantaneous value of X.
- The process described by eq. (9.127) must be compatible with the equilibrium distribution (9.124)

Given these hypotheses, the random force $f(t)$ satisfies the relations

$$\langle f(t) \rangle = 0, \qquad \langle f(t) f(t') \rangle = 2\Lambda\, \delta(t - t'), \tag{9.128}$$

in which Λ satisfies a relation analogous to Einstein's, which we will determine further on. By following the line of reasoning of sec. 9.5, it is easy to see that the probability

distribution $P(x,t)$ satisfies the Fokker-Planck equation:

$$\frac{\partial}{\partial t}P(x,t)=\frac{\partial}{\partial x}\left\{-\left[b(x)\,P+\Lambda\frac{\partial P}{\partial x}\right]\right\}. \tag{9.129}$$

By imposing the stationarity of the equilibrium distribution $P_{\rm eq}(x)$, we obtain the expression of the systematic velocity $b(x)$:

$$b(x)=-\frac{\Lambda}{k_{\rm B}T}\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{F}(x). \tag{9.130}$$

This has the form of a phenomenological equation of the type $b=\Gamma F$, where $F=-\mathrm{d}\mathcal{F}/\mathrm{d}x$ is a force, and $\Gamma=\Lambda/k_{\rm B}T$ is a kinetic coefficient (analogous to the damping coefficient in fluids). One usually assumes that Γ depends less strongly than Λ on temperature, and one therefore writes the Langevin equation in the form

$$\frac{\mathrm{d}x}{\mathrm{d}t}=-\Gamma\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}x}+f(t), \tag{9.131}$$

with the following conditions for the random force $f(t)$:

$$\big\langle f(t)\big\rangle=0,\qquad \big\langle f(t)f(t')\big\rangle=2k_{\rm B}T\,\Gamma\,\delta(t-t'). \tag{9.132}$$

By means of these essentially phenomenological considerations, we have obtained a nonlinear Langevin equation satisfied by the fluctuations of X. Recall that the availability $\mathcal{F}(x)$ reaches its minimum for the equilibrium value $x_{\rm eq}$ of X, and that both X and $\mathcal{F}$ are extensive quantities with respect to the system's size (which we denote by N). Therefore,

$$\mathcal{F}(x)=\mathcal{F}(x_{\rm eq})+\frac{1}{2}\left.\frac{\mathrm{d}^2\mathcal{F}}{\mathrm{d}x^2}\right|_{x=x_{\rm eq}}(x-x_{\rm eq})^2+\cdots. \tag{9.133}$$

Since $P_{\rm eq}(x)\propto \mathrm{e}^{-\mathcal{F}(x)/k_{\rm B}T}$ and $\mathrm{d}^2\mathcal{F}/\mathrm{d}x^2\sim N^{-1}$, the typical values of $\left|x-x_{\rm eq}\right|$ are of the order of $N^{1/2}$, and further terms in this expansion are usually negligible in the thermodynamic limit.

In this approximation, the equation satisfied by X is linear. By introducing the deviation $\delta x=x-x_{\rm eq}$ we obtain

$$\frac{\mathrm{d}\,\delta x}{\mathrm{d}t}=-\frac{\Gamma}{\chi}\,\delta x+f(t), \tag{9.134}$$

where we have introduced the susceptibility χ, defined by

$$\frac{1}{\chi}=\left.\frac{\mathrm{d}^2\mathcal{F}(x)}{\mathrm{d}x^2}\right|_{x=x_{\rm eq}}. \tag{9.135}$$

It is easy to see that χ is actually a susceptibility. Let us introduce a perturbation of the form $-\lambda X(\nu)$ into the Hamiltonian $\mathcal{H}(\nu)$ and calculate the change in the equilibrium value of X. We then obtain

$$\left.\frac{\partial\,\langle X\rangle}{\partial\lambda}\right|_{\lambda=0}=\chi. \tag{9.136}$$

Near the critical point, the order parameter susceptibility χ diverges. In this case, it is necessary to take the other terms in the expansion (9.133) into account.

From this equation, retracing the calculations of sec. 5.14, we obtain the correlation function of X;

$$C(t) = \langle X(t)X(0)\rangle = k_B T\, \chi\, \exp\left(-\frac{\Gamma}{\chi}\,|t|\right), \tag{9.137}$$

and the corresponding Fourier transform

$$C(\omega) = \int \mathrm{d}t\, \mathrm{e}^{\mathrm{i}\omega t}\, C(t) = \frac{2\Gamma\, k_B T}{\omega^2 + (\Gamma/\chi)^2}. \tag{9.138}$$

The characteristic relaxation time τ_r is given by χ/Γ. It diverges approaching the critical point, if X is the order parameter. On the one hand, this guarantees that the approximations made over the correlations of the random forces become more acceptable, while on the other hand, it makes validity of the expansion (9.133) more doubtful.

If we continue, however, to describe the system in the context of the linearized Langevin equation, we obtain $\tau_r \propto \chi \propto |T - T_c|^{-\gamma}$. This is the *classical* result, due to van Hove [Hove54]. The divergence of the relaxation times in proximity to the critical point is a general phenomenon and is known as **critical slowing down**. A more detailed description of the dynamics close to a critical point can be obtained by renormalization-group methods, thoroughly described, for example, in [Taub14].

9.11 Response Functions

The argument of sec. 9.10 can be generalized to the case in which, instead of a single variable X, we are dealing with a collection $X = (X_1, \ldots, X_r)$ of dynamic variables. These variables are defined as functions of the microscopic state x, which is itself identified by the collection $(\boldsymbol{p}^N, \boldsymbol{r}^N)$ of the momenta and positions of the particles that make up the system. The equilibrium probability distribution is expressed as a function of the availability $\mathcal{F}(X) = \mathcal{F}(X_1, \ldots, X_r)$. The Langevin equation for the instantaneous value $X_\alpha(t) = X_\alpha(x(t))$ of the observable X_α can be written

$$\frac{\mathrm{d}X_\alpha}{\mathrm{d}t} = b_\alpha(X(t)) + f_\alpha(t), \qquad \alpha \in \{1, \ldots, r\}. \tag{9.139}$$

To proceed, we must specify how the observables X_α behave under time inversion. The time-reversal operator I is defined in the space of the microstates x. Given $x = ((\boldsymbol{p}_1, \boldsymbol{r}_1), \ldots, (\boldsymbol{p}_N, \boldsymbol{r}_N))$, one has

$$x' = \mathsf{I}\, x = ((-\boldsymbol{p}_1, \boldsymbol{r}_1), \ldots, (-\boldsymbol{p}_N, \boldsymbol{r}_N)). \tag{9.140}$$

If magnetic fields and Coriolis forces are absent, the Hamiltonian is invariant under time reversal.[3] Then, if the Hamiltonian does not depend explicitly on time, given a solution $x(t)$

[3] We also need to disregard the violation of time reversal characteristic of weak interactions.

of the canonical equations of motion that satisfies the initial condition $x(0) = x_0$, its time-reversed image $\widehat{x}(t) = \mathsf{I}\, x(-t)$ is a solution of the equations of motion that satisfies the initial condition $\widehat{x}(0) = \mathsf{I}\, x_0$.

We can then define the observables $X_\alpha(x)$ in such a way that they are eigenfunctions of I:

$$\mathsf{I} X_\alpha(x) = X_\alpha(\mathsf{I}\, x) = \epsilon_\alpha X_\alpha(x), \tag{9.141}$$

where $\epsilon_\alpha = \pm 1$ is the **parity** of X_α under time inversion.

Let us now calculate the effect of time reversal on the correlation function $\langle X_\alpha(t)\, X_\beta(0)\rangle$. We fix a state x_0 at $t = 0$. We then shift to its time-reversed image $\widehat{x}_0 = \mathsf{I}\, x_0$ and let the system evolve. After time t, the system will be in the state $\mathsf{I}\, x(-t)$. The corresponding value of X_α is equal to $\mathsf{I} X_\alpha(-t) = \epsilon_\alpha X_\alpha(-t)$. On the other hand, we have $X_\beta(\widehat{x}_0) = \epsilon_\beta X_\beta(x_0)$ by definition, and, since the Hamiltonian is invariant, the probabilities of x_0 and of $\widehat{x}_0$ are equal at equilibrium. We conclude that, at equilibrium, we have

$$\langle X_\alpha(t)\, X_\beta(0)\rangle = \epsilon_\alpha \epsilon_\beta \langle X_\alpha(-t)\, X_\beta(0)\rangle. \tag{9.142}$$

Therefore, the correlation functions of quantities that have the same parity with respect to time reversal are *even* functions of time, whereas the others are *odd*.

For simplicity we now assume that the observables X_α are all even under time reversal and that the correlation of random forces does not depend on the values of the observables. We can then define the observables so that the correlation matrix of the random forces is diagonal:

$$\langle f_\alpha(t) f_\beta(t')\rangle = 2 k_\mathrm{B} T\, \Gamma_\alpha\, \delta_{\alpha\beta}\, \delta(t - t'), \qquad \alpha, \beta \in \{1, \ldots, r\}. \tag{9.143}$$

Then the Fokker-Planck equation assumes the form

$$\frac{\partial}{\partial t} P(x, t) = \sum_\alpha \frac{\partial}{\partial X_\alpha} \left[-b_\alpha(X)\, P + k_\mathrm{B} T\, \Gamma_\alpha \frac{\partial P}{\partial X_\alpha} \right]. \tag{9.144}$$

Imposing that $P_\mathrm{eq}(X) \propto \exp\left(-\mathcal{F}(X)/k_\mathrm{B} T\right)$ is a time-independent solution of the Fokker-Planck equation, we obtain the expression of the velocities $b_\alpha(X)$:

$$b_\alpha(X) = -\Gamma_\alpha \frac{\partial \mathcal{F}}{\partial X_\alpha}, \qquad \alpha \in \{1, \ldots, r\}. \tag{9.145}$$

With this definition, the Fokker-Planck equation assumes the form

$$\frac{\partial}{\partial t} P(X, t) = \sum_\alpha \Gamma_\alpha \frac{\partial}{\partial X_\alpha} \left[\frac{\partial \mathcal{F}}{\partial X_\alpha}\, P + k_\mathrm{B} T \frac{\partial P}{\partial X_\alpha} \right]. \tag{9.146}$$

The Fokker-Planck equation assumes a similar form when the correlation matrix of the random force is not diagonal:

$$\frac{\partial P}{\partial t} = \sum_{\alpha\beta} \Gamma_{\alpha\beta} \frac{\partial}{\partial X_\alpha} \left[\frac{\partial \mathcal{F}}{\partial X_\beta} P + k_\mathrm{B} T \frac{\partial P}{\partial X_\beta} \right]. \tag{9.147}$$

Let us now suppose that we perturb our system, initially (at time $t = 0$) at thermodynamic equilibrium, by adding a term of the form $-\sum_\alpha \lambda_\alpha(t) X_\alpha(x)$ to the Hamiltonian. This perturbation reverberates in change of the availability $\mathcal{F}$, which comes to depend on time in the form

$$\mathcal{F}(X, t) = \mathcal{F}_0(X) - \sum_\alpha \lambda_\alpha(t) X_\alpha, \tag{9.148}$$

where $\mathcal{F}_0$ is the "unperturbed" expression of the availability, that is, the one that we have considered so far. In order to persuade ourselves of this result, it is sufficient to imagine performing the calculation of the integral (9.125) in the presence of the external field λ, but keeping the value of the X_α variables constrained. One can then take the contribution due to the field out of the integral, and one obtains expression (9.148).

Because of the perturbation, the average values of the observables X_α will depend on time. Let us redefine the observables in such a fashion that their average equilibrium values vanish. We want to calculate their average values at instant $t > 0$, to the first order in the perturbation $\lambda_\alpha(t)$. We have in general

$$\left\langle X_\alpha(t) \right\rangle_\lambda = \int \mathrm{d}t' \sum_\beta \chi_{\alpha\beta}(t, t')\, \lambda_\beta(t') + \mathrm{O}\left(\lambda^2\right), \tag{9.149}$$

where

$$\left\langle X_\alpha(t) \right\rangle_\lambda = \int \mathrm{d}X'\, X'_\alpha\, P_\lambda(X', t), \tag{9.150}$$

in which $P_\lambda(X, t)$ is the solution of the Fokker-Planck equation (9.146) with the perturbed availability (9.148) satisfying the initial condition $P(X, 0) = P_{\mathrm{eq}}(X)$. Equation (9.149) defines the matrix $\chi(t, t') = (\chi_{\alpha\beta}(t, t'))$ of the **response functions**.

The response functions $\chi_{\alpha\beta}(t, t')$ have two important properties:

- We have $\chi_{\alpha\beta}(t, t') = 0$ for $t' > t$, because effects cannot precede causes. This property is called **causality**.
- Since in the absence of a perturbation the system starts at equilibrium, the response functions satisfy time-translation invariance:

  $$\chi_{\alpha\beta}(t, t') = \chi_{\alpha\beta}(t + \tau, t' + \tau), \qquad \forall \tau. \tag{9.151}$$

  We can therefore write $\chi_{\alpha\beta}(t, t') = \chi_{\alpha\beta}(t - t', 0)$, and we can drop the second argument. Because of causality, we have

  $$\chi_{\alpha\beta}(t) = 0, \qquad \text{for } t < 0. \tag{9.152}$$

We now show that the response functions can be expressed in terms of an equilibrium correlation function. For simplicity, let us consider a single observable X that satisfies the Langevin equation

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -\Gamma \frac{\partial \mathcal{F}}{\partial X} + f(t), \tag{9.153}$$

where the Gaussian random force $f(t)$ satisfies the usual conditions:

$$\langle f(t)\rangle = 0, \qquad \langle f(t)f(t')\rangle = 2\Gamma k_B T\delta(t-t'), \quad \forall t, t'. \tag{9.154}$$

We further define

$$C(t-t') = \langle X(t_2)X(t_1)\rangle_{eq} = \int dX_1 \int dX_2\, X_1\, X_2\, P(X_2, t_2|X_1, t_1)\, P_{eq}(X_1), \tag{9.155}$$

where $P(X_2, t_2|X_1, t_1)$ is the probability that X has the value X_2 at time t_2, given that it had the value X_1 at time t_1, while evolving with the unperturbed availability $\mathcal{F}_0(X)$, and where $P_{eq}(X) \propto e^{-\mathcal{F}_0(X)/k_B T}$ is the unperturbed equilibrium distribution. We recall that we have defined X such that $\langle X(t)\rangle_{eq} = 0$.

To evaluate $\chi(t)$, we introduce a perturbation $\lambda(t)$ of the form

$$\lambda(t) = \theta_H(-t)\, e^{\varepsilon t}\, \lambda_0, \tag{9.156}$$

where $\theta_H(t)$ is Heaviside's step function, and ε is a small positive quantity. This perturbation grows very slowly from 0 (for $t = -\infty$) to λ_0 (for $t = 0$), and is then suddenly switched off. We then evaluate $\langle X(t)\rangle_\lambda$ for $t > 0$. By the definition of the response function, we obtain

$$\langle X(t)\rangle_\lambda = \lambda_0 \int_{-\infty}^{0} dt'\, \chi(t-t'), \tag{9.157}$$

where we have implicitly taken the $\varepsilon \to 0$ limit. On the other hand, we have

$$\langle X(t)\rangle_\lambda = \int dX \int dX_0\, X_0\, P(X, t|X_0, 0)\, P_\lambda(X_0, 0), \tag{9.158}$$

where $P_\lambda(X, 0)$ is the probability distribution of X at $t = 0$ under the perturbation. But because of the form of the perturbation, this distribution is the equilibrium distribution of X with the perturbed availability $\mathcal{F}(X) = \mathcal{F}_0(X) - \lambda_0 x$. We have therefore

$$P_\lambda(X, 0) = \exp\left(\frac{F(\lambda_0) - \mathcal{F}_0(X) + \lambda_0 X}{k_B T}\right), \tag{9.159}$$

where

$$F(\lambda) = -k_B T \int dX\, \exp\left(-\frac{\mathcal{F}_0(X) - \lambda X}{k_B T}\right). \tag{9.160}$$

Expanding the expression (9.159) to first order in λ_0, and taking into account the fact that $dF/d\lambda|_{\lambda=0} = -\langle X\rangle_{eq} = 0$, we obtain

$$P_\lambda(X, 0) \simeq \frac{\lambda_0 X}{k_B T}\, P_{eq}(X). \tag{9.161}$$

Therefore

$$\begin{aligned}\langle X(t)\rangle_\lambda &\simeq \frac{\lambda_0}{k_B T} \int dX \int dX_0\, X\, X_0\, P(X, t|X_0, 0)\, P_{eq}(X_0) \\ &= \frac{\lambda_0}{k_B T} \langle X(t)X(0)\rangle = \frac{\lambda_0}{k_B T} C(t).\end{aligned} \tag{9.162}$$

This result is called **Onsager's regression hypothesis**. Lars Onsager *postulated* it in 1931 [Onsa31a, Onsa31b], arguing that the system cannot "know" whether a nonvanishing value of the observable X is due to a spontaneous fluctuation or to a perturbation that occurred in the past—the regression of its average toward its equilibrium value must be proportional to the regression of the correlation function toward zero.

Comparing this result with eq. (9.157), we obtain the relation

$$C(t) = k_B T \int_{-\infty}^{0} \mathrm{d}t' \, \chi(t - t') = k_B T \int_{t}^{\infty} \mathrm{d}\tau \, \chi(\tau). \tag{9.163}$$

Taking the derivative of this result with respect to t, we finally obtain

$$\chi(t) = -\frac{1}{k_B T} \frac{\mathrm{d}}{\mathrm{d}t} C(t), \qquad t > 0. \tag{9.164}$$

This result can be easily generalized to several observables X_α. Taking into account causality, we obtain the general expression

$$\chi_{\alpha\beta}(t) = -\frac{\theta_H(t)}{k_B T} \frac{\mathrm{d}}{\mathrm{d}t} C_{\alpha\beta}(t), \tag{9.165}$$

where

$$C_{\alpha\beta}(t) = \left\langle X_\alpha(t) X_\beta(0) \right\rangle_{\mathrm{eq}}. \tag{9.166}$$

We derived this equation in the context of the Langevin equation—it is possible, however, to see that the result holds also in the case in which the evolution of the system is described by the canonical equations of motion (and therefore the probability distribution obeys the Liouville equation).

Time-reversal invariance has an important consequence on the symmetry of the matrix of response functions. For $t > 0$, we obtain in fact

$$\begin{aligned} \chi_{\alpha\beta}(t) &= -\frac{1}{k_B T} \frac{\mathrm{d}}{\mathrm{d}t} \left\langle X_\alpha(t) X_\beta(0) \right\rangle = -\frac{1}{k_B T} \epsilon_\alpha \epsilon_\beta \frac{\mathrm{d}}{\mathrm{d}t} \left\langle X_\alpha(-t) X_\beta(0) \right\rangle \\ &= -\frac{1}{k_B T} \epsilon_\alpha \epsilon_\beta \frac{\mathrm{d}}{\mathrm{d}t} \left\langle X_\alpha(0) X_\beta(t) \right\rangle = \epsilon_\alpha \epsilon_\beta \, \chi_{\beta\alpha}(t). \end{aligned} \tag{9.167}$$

This relation is at the basis of the **Onsager relations**, which express the symmetry of the kinetic coefficients. We will discuss these relations further on.

9.12 Fluctuation-Dissipation Theorem

Let us return to eq. (9.165), which we can write in the form

$$\chi_{\alpha\beta}(t) = \frac{\theta_H(t)}{k_B T} \left(-\frac{\mathrm{d}}{\mathrm{d}t} \right) C_{\alpha\beta}(\omega). \tag{9.168}$$

Taking the Fourier transform of both sides, and recalling that the Fourier transform of a product of two functions is given by the convolution of the transforms (cf. sec. D.2),

we obtain

$$\begin{aligned}\chi_{\alpha\beta}(\omega) &= \int_{-\infty}^{+\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t}\, \chi_{\alpha\beta}(t) \\ &= -\frac{1}{k_{\mathrm{B}}T} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega'}{2\pi}\, \theta_{\mathrm{H}}(\omega-\omega')\, (\mathrm{i}\omega')\, C_{\alpha\beta}(\omega'),\end{aligned} \tag{9.169}$$

where $\theta_{\mathrm{H}}(\omega)$ is the Fourier transform of the Heaviside step function.

To evaluate $\theta_{\mathrm{H}}(\omega)$, we observe that the integral that defines it does not converge, unless one adds an exponentially decreasing factor. We thus define

$$\theta_{\mathrm{H}}(\omega) = \lim_{\varepsilon\to 0^+} \int_0^{\infty} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}\omega t-\varepsilon t} = \lim_{\varepsilon\to 0^+} \frac{1}{\mathrm{i}\omega+\varepsilon}. \tag{9.170}$$

One has

$$\frac{1}{\mathrm{i}\omega+\varepsilon} = \frac{-\mathrm{i}\omega+\varepsilon}{\omega^2+\varepsilon}. \tag{9.171}$$

Let us consider the integral of this expression multiplied by an arbitrary regular real-valued function $\phi(\omega)$:

$$\begin{aligned}\lim_{\varepsilon\to 0^+} \int \mathrm{d}\omega\, \frac{1}{\mathrm{i}\omega+\varepsilon}\phi(\omega) = \lim_{\varepsilon\to 0^+} \Bigg[&-\mathrm{i}\int_{-\infty}^{+\infty} \mathrm{d}\omega\, \frac{\omega}{\omega^2+\varepsilon^2}\phi(\omega) \\ &+ \int_{-\infty}^{+\infty} \mathrm{d}\omega\, \frac{\varepsilon}{\omega^2+\varepsilon^2}\phi(\omega)\Bigg].\end{aligned} \tag{9.172}$$

Let us separately consider the real and imaginary parts of this expression. We have for the real part

$$\begin{aligned}\lim_{\varepsilon\to 0^+} \int_{-\infty}^{+\infty} \mathrm{d}\omega\, \frac{\varepsilon}{\omega^2+\varepsilon^2}\phi(\omega) &= \lim_{\varepsilon\to 0^+} \int_{-\infty}^{+\infty} \mathrm{d}\omega'\, \frac{\varepsilon^2}{(\varepsilon\omega')^2+\varepsilon^2}\phi(\varepsilon\omega') \\ &= \pi\, \phi(0).\end{aligned} \tag{9.173}$$

For the imaginary part, notice that $\omega/(\omega^2+\varepsilon^2)$ yields ω^{-1} for $|\omega| \gg \varepsilon$, but yields ω/ε^2 for $|\omega| \ll \varepsilon$. Since $\phi(\omega)$ is regular around $\omega = 0$, the contribution to the integral for $|\omega| \lesssim \varepsilon$ vanishes in the limit. We obtain therefore

$$\lim_{\varepsilon\to 0^+} \int_{-\infty}^{+\infty} \mathrm{d}\omega\, \frac{\omega}{\omega^2+\varepsilon^2}\phi(\omega) = \lim_{\varepsilon\to 0^+} \left[\int_{-\infty}^{-\varepsilon} \mathrm{d}\omega\, \frac{1}{\omega}\phi(\omega) + \int_{+\varepsilon}^{+\infty} \mathrm{d}\omega\, \frac{1}{\omega}\phi(\omega)\right]. \tag{9.174}$$

The expression on the right-hand side is known as the **principal value** of the singular integral $\int \mathrm{d}\omega\, \phi(\omega)/\omega$, and we will denote it by $\mathcal{P}\int \mathrm{d}\omega\, \phi(\omega)/\omega$. We can summarize these results by setting

$$\lim_{\varepsilon\to 0^+} \frac{1}{\mathrm{i}\omega+\varepsilon} = -\mathrm{i}\mathcal{P}\frac{1}{\omega} + \pi\,\delta(\omega). \tag{9.175}$$

Substituting in eq. (9.169) we obtain

$$\chi_{\alpha\beta}(\omega) = \frac{1}{k_B T}\left[\mathcal{P}\int_{-\infty}^{+\infty}\frac{d\omega'}{2\pi}\,\frac{\omega'}{\omega-\omega'}C_{\alpha\beta}(\omega') - \frac{i\omega}{2}C_{\alpha\beta}(\omega)\right]. \tag{9.176}$$

If the $C_{\alpha\beta}(t)$ function is even, its Fourier transform is real. This expression therefore allows one to express a relation between the real and the imaginary parts of the response function, called the **Kramers-Kronig relation**, which expresses the analyticity of $\chi(\omega)$ (with our conventions) in the lower half-plane of the complex plane. Setting

$$\chi_{\alpha\beta}(\omega) = \chi'_{\alpha\beta}(\omega) + i\chi''_{\alpha\beta}(\omega), \tag{9.177}$$

with χ' and χ'' real, we have

$$\chi'_{\alpha\beta}(\omega) = -\mathcal{P}\int_{-\infty}^{+\infty}\frac{d\omega'}{\pi}\,\frac{1}{\omega-\omega'}\,\chi''_{\alpha\beta}(\omega'). \tag{9.178}$$

(If the correlation function is odd, its transform is purely imaginary, and one has an analogous relation in which the real and imaginary parts are switched.)

To close these considerations, let us obtain the relation between the correlation and the response in Fourier space:

$$\chi''_{\alpha\beta}(\omega) = -\frac{\omega}{k_B T}\,C_{\alpha\beta}(\omega). \tag{9.179}$$

This is the most frequently used form of the relation between fluctuation and response, and is called the **fluctuation-dissipation theorem**. Indeed, the imaginary part $\chi''(\omega)$ of the response function is related to the dissipation of the external work performed by the perturbation.

Let us show this in a simple case in which we are dealing with a single observable X and a sinusoidal perturbation of frequency ω:

$$\lambda(t) = \mathrm{Re}\left(\lambda_0\, e^{-i\omega t}\right) = \frac{1}{2}\left(\lambda_0 e^{-i\omega t} + \lambda_0^* e^{i\omega t}\right). \tag{9.180}$$

We want to evaluate the time average of the mean dissipated power, given by

$$\begin{aligned}\overline{W} &= \lim_{t\to\infty}\frac{1}{t}\int_0^t dt'\,\lambda(t)\left\langle\frac{dX(t')}{dt'}\right\rangle = \lim_{t\to\infty}\frac{1}{t}\int_0^t dt'\,\left(-\langle X(t')\rangle\right)\frac{d\lambda(t')}{dt'}\\ &= \lim_{t\to\infty}\frac{1}{t}\int_0^t dt'\,\langle X(t')\rangle\,\frac{i\omega}{2}\left(\lambda_0 e^{-i\omega t'} - \lambda_0^* e^{i\omega t'}\right),\end{aligned} \tag{9.181}$$

where we have integrated by parts and neglected the boundary terms. On the other hand, we have

$$\begin{aligned}\langle X(t')\rangle &= \int_{-\infty}^{+\infty} dt''\,\chi(t'-t'')\,\lambda(t'') = \int_{-\infty}^{+\infty} dt''\,\chi(t'')\lambda(t'-t'')\\ &= \int_{-\infty}^{+\infty} dt''\,\frac{\chi(t'')}{2}\left(\lambda_0 e^{-i\omega(t'-t'')} + \lambda_0^* e^{i\omega(t'-t'')}\right).\end{aligned} \tag{9.182}$$

Substituting this expression in eq. (9.181), we obtain

$$\begin{aligned}\overline{W} &= \lim_{t\to\infty}\frac{1}{t}\int_0^t \mathrm{d}t' \int_{-\infty}^{+\infty}\mathrm{d}t''\,\frac{\mathrm{i}\omega\chi(t'')}{4}\left(\lambda_0 \mathrm{e}^{-\mathrm{i}\omega(t'-t'')}+\lambda_0^* \mathrm{e}^{\mathrm{i}\omega(t'-t'')}\right)\\ &\quad\times\left(\lambda_0 \mathrm{e}^{-\mathrm{i}\omega t'}-\lambda_0^* \mathrm{e}^{\mathrm{i}\omega t'}\right)\\ &= \int_{-\infty}^{+\infty}\mathrm{d}t''\,\frac{\mathrm{i}\omega\chi(t'')}{4}\lim_{t\to\infty}\frac{1}{t}\int_0^t \mathrm{d}t'\\ &\quad\times\left[\lambda_0^2 \mathrm{e}^{-\mathrm{i}\omega(2t'-t'')}-\lambda_0^{*2}\mathrm{e}^{\mathrm{i}\omega(2t'-t'')}+|\lambda_0|^2\left(\mathrm{e}^{-\mathrm{i}\omega t''}-\mathrm{e}^{\mathrm{i}\omega t''}\right)\right].\end{aligned} \tag{9.183}$$

When integrated over t', the first two terms remain bounded and their contribution thus vanishes as $1/t$. The third term is independent of t' and therefore its integral cancels the factor $1/t$. The difference of the exponentials yields a sine. We thus obtain

$$\overline{W} = \frac{\omega|\lambda_0|^2}{2}\int_{-\infty}^{+\infty}\mathrm{d}t\,\chi(t)\,\sin(\omega t) = \frac{\omega|\lambda_0|^2}{2}\,\mathrm{Im}\,\chi(\omega) = \frac{\omega|\lambda_0|^2}{2}\chi''(\omega). \tag{9.184}$$

Since the dissipated power is necessarily positive, it follows that $\omega\chi''(\omega) > 0$.

Exercise 9.5 (Brownian motion of a harmonic oscillator) Consider a Brownian particle of mass m subject to a harmonic potential of Hooke's constant $m\omega_0^2$ in one dimension, immersed in a fluid at the temperature T, with a damping coefficient equal to γ.

1. Write down the evolution equation of the system when the system is subjected to a time-dependent applied force $f(t)$.
2. Assuming that the system was in equilibrium in the distant past, express the average instantaneous position $\langle x(t)\rangle$ in terms of the response function $\chi_{xx}(t)$ and $f(t)$. By solving the Langevin equation, give the expressions of the response function. Deduce the form of the response function $\chi_{xv}(t)$ that expresses $\langle v(t)\rangle$ as a functional of $f(t)$.
3. Evaluate the correlation function $C_{xx}(t) = \langle x(t)x(0)\rangle$, assuming that $\langle x^2(0)\rangle$ is equal to the equilibrium value.
4. Evaluate the Fourier transform $C_{xx}(\omega) = \int_{-\infty}^{+\infty}\mathrm{d}t\,\mathrm{e}^{\mathrm{i}\omega t}\,C_{xx}(t)$ and compare it with the imaginary part $\chi''_{xx}(\omega)$ of the Fourier transform of the corresponding response function $\chi_{xx}(t)$.

Exercise 9.6 Consider a Brownian particle in contact with a heat reservoir. We denote by x the coordinate of the particle and by $(y_1, \ldots, y_N)$ the coordinates of the reservoir. The Hamiltonian of the system is given by

$$\mathcal{H} = \mathcal{H}_0(x) - xf + \mathcal{H}_\mathrm{b}(y_1, \ldots, y_N).$$

We assume

$$f = \sum_i c_i y_i,$$

and

$$\mathcal{H}_0 = \frac{m\dot{x}^2}{2} + V(x).$$

We also assume that $\mathcal{H}_\mathrm{b}$ is the Hamiltonian for a system of classical oscillators. This implies the exact validity of linear-response theory, yielding

$$f(t) = f_\mathrm{b}(t) + \int_{-\infty}^{+\infty} \mathrm{d}t'\, \chi_\mathrm{b}(t-t')x(t'),$$

where

$$\chi_\mathrm{b}(t) = \begin{cases} -\beta \mathrm{d}C_\mathrm{b}(t)/\mathrm{d}t, & \text{for } t > 0; \\ 0, & \text{otherwise.} \end{cases}$$

Here, $C_\mathrm{b}(t)$ is the correlation function of the isolated reservoir:

$$C_\mathrm{b}(t) = \left\langle \delta f(t)\, \delta f(0) \right\rangle_\mathrm{b}.$$

Show that $x(t)$ satisfies the equation

$$m\ddot{x}(t) = \tilde{f}[x(t)] + \delta f(t) - \beta \int_0^t \mathrm{d}t'\, C_{\mathrm{b}(t-t')}\dot{x}(t'),$$

where

$$\tilde{f}(x) = -\frac{\mathrm{d}\tilde{V}}{\mathrm{d}x};$$

$$\tilde{V}(x) = V(x) - \beta C(0)\frac{x^2}{2};$$

$$\delta f(t) = f_\mathrm{b}(t) - \beta C_\mathrm{b}(t)x(0).$$

Thus $\delta f(t)$ is a Gaussian force with vanishing average and correlation equal to $C_\mathrm{b}(t-t')$.

9.13 Onsager Reciprocity Relations

Let us assume that X_α, $\alpha = 1, \ldots, r$, are observables, defined in such a manner that their average value at equilibrium vanishes:

$$\langle X_\alpha \rangle = 0, \qquad \forall \alpha. \tag{9.185}$$

We will always denote the equilibrium average by $\langle \ldots \rangle$. By the relation between response and correlation, we can evaluate the nonequilibrium average of the observable X_α at instant t, under a perturbation of type $-\sum_\alpha h_\alpha(t)X_\alpha$ to the Hamiltonian. As we have seen, we obtain

$$\langle X_\alpha \rangle_t = \int_{-\infty}^{t} \mathrm{d}t' \sum_\ell \chi_{\alpha\beta}(t-t')\, h_\beta(t'), \tag{9.186}$$

where

$$\chi_{\alpha\beta}(t-t') = -\frac{\theta_{\mathrm{H}}(t-t')}{k_{\mathrm{B}}T}\frac{\mathrm{d}}{\mathrm{d}t}\langle X_\alpha(t)X_\beta(t')\rangle. \tag{9.187}$$

Let us assume that the perturbation is switched off, and that we are interested in the decay to zero of the averages. In a linear regime, it is reasonable to assume that the nonequilibrium averages $\langle X_k\rangle_t$ satisfy a linear differential equation of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X_\alpha\rangle_t = -\sum_\beta M_{\alpha\beta}\langle X_\beta\rangle_t. \tag{9.188}$$

In reality, this form is valid only if one operates a coarse-graining over time—in other words, if one gives up observing the dynamics over time intervals that are too small (on the order of the correlation time of the noise).

Onsager's regression hypothesis (9.162) implies that the correlation function satisfies an analogous differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\langle X_\alpha(t)X_\beta(0)\right\rangle = -\sum_\gamma M_{\alpha\gamma}\left\langle X_\gamma(t)X_\beta(0)\right\rangle, \qquad \forall k, \quad t>0. \tag{9.189}$$

If X_α and X_β are both time-reversal-invariant, and if the Hamiltonian is also invariant, one has

$$\left\langle X_\alpha(t)X_\beta(0)\right\rangle = \left\langle X_\beta(t)X_\alpha(0)\right\rangle. \tag{9.190}$$

By taking the derivative of both sides of this equation with respect to t and substituting eq. (9.189), we obtain

$$\sum_\gamma M_{\alpha\gamma}\left\langle X_\gamma(t)X_\beta(0)\right\rangle = \sum_\gamma M_{\beta\gamma}\left\langle X_\gamma(t)X_\alpha(0)\right\rangle. \tag{9.191}$$

By setting $t=0$ in this relation, we obtain

$$\sum_\gamma M_{\alpha\gamma}\left\langle X_\gamma X_\beta\right\rangle = \sum_\gamma M_{\beta\gamma}\left\langle X_\gamma X_\alpha\right\rangle, \tag{9.192}$$

where

$$\left\langle X_\alpha X_\beta\right\rangle = \left\langle X_\alpha(0)X_\beta(0)\right\rangle \tag{9.193}$$

is the instantaneous correlation at equilibrium of the observables X_α and X_β.

By introducing the matrix $\mathsf{L}=(L_{\alpha\beta})$, where

$$L_{\alpha\beta} = \frac{1}{k_{\mathrm{B}}}\sum_\gamma M_{\alpha\gamma}\left\langle X_\gamma X_\beta\right\rangle, \tag{9.194}$$

this result can be written

$$L_{\alpha\beta} = L_{\beta\alpha}, \tag{9.195}$$

in other words, $\mathsf{L}^{\mathsf{T}} = \mathsf{L}$, which implies that the matrix L is *symmetrical*.

These relations, obtained by L. Onsager in 1931 [Onsa31a, Onsa31b], are of great importance, as regards both fundamentals and applications, to the extent that they earned their author the 1968 Nobel Prize for chemistry. They are called the **Onsager reciprocity relations**.

9.14 Affinities and Fluxes

The importance of the Onsager reciprocity relations can be better understood by considering how a system responds to a perturbation that takes it out of thermodynamic equilibrium. In order to clarify our ideas, we consider a system in contact with a reservoir. The whole (system and reservoir) is isolated from the environment.

We denote the extensive thermodynamic observables by $X = (X_1, \ldots, X_r)$ for both the system and the reservoir. The thermodynamic entropy of the whole is maximal at equilibrium. When the system is out of equilibrium it will be in general possible to exchange some amount of one (or more) of these observables between system and reservoir, in order to increase the total entropy. This implies that for some α we have

$$A_\alpha = \frac{\partial S}{\partial X_\alpha} - \frac{\partial S^{(\mathrm{r})}}{\partial X_\alpha^{(\mathrm{r})}} \neq 0. \tag{9.196}$$

Here S is the entropy of the system, $S^{(\mathrm{r})}$ that of the reservoir, and analogously for the observables X. We assume that the thermodynamical state of the system and the reservoir is identified by the values of the observables X. Therefore the equilibrium state corresponds to $A_\alpha = 0$, $\forall\alpha$. We can thus consider the A_α as the *forces* that take the system out of equilibrium. These generalized forces are often called **affinities**.

For example, if X is the internal energy E, the corresponding affinity is given by

$$A_E = \frac{1}{T} - \frac{1}{T^{(\mathrm{r})}}. \tag{9.197}$$

On the other hand, if X is the volume V, one instead has

$$A_V = \frac{p}{T} - \frac{p^{(\mathrm{r})}}{T^{(\mathrm{r})}}. \tag{9.198}$$

If the observable is the number of moles of the chemical species k, one has

$$A_k = -\frac{\mu_k}{T} + \frac{\mu_k^{(\mathrm{r})}}{T^{(\mathrm{r})}}, \tag{9.199}$$

where μ_k is the corresponding chemical potential.

If the system is no more in thermodynamic equilibrium, the values of the observables will start to change with time. We can define the **fluxes** J_α as the time derivatives

of the X_α:

$$J_\alpha = \frac{\mathrm{d}X_\alpha}{\mathrm{d}t}. \tag{9.200}$$

This change shows up as a change in the value of the total entropy $S^{\mathrm{tot}} = S + S^{(\mathrm{r})}$:

$$\frac{\mathrm{d}S^{\mathrm{tot}}}{\mathrm{d}t} = \sum_\alpha \left(\frac{\partial S}{\partial X_\alpha} - \frac{\partial S^{(\mathrm{r})}}{\partial X_\alpha^{(\mathrm{r})}} \right) \frac{\mathrm{d}X_\alpha}{\mathrm{d}t} = \sum_\alpha A_\alpha J_\alpha. \tag{9.201}$$

The entropy production is therefore given by the sum of the products of the affinities times the corresponding fluxes. The second principle of thermodynamics guarantees that since the whole is isolated, total entropy cannot but increase. One therefore has

$$\frac{\mathrm{d}S^{\mathrm{tot}}}{\mathrm{d}t} = \sum_\alpha A_\alpha J_\alpha \geq 0. \tag{9.202}$$

In a generic situation of nonequilibrium, it is in general possible to express the fluxes J_α as functions of the affinities A_β and of the equilibrium parameters of the system. One will therefore have $J_\alpha = J_\alpha(F)$, $\alpha = 1, \ldots, r$. The fluxes vanish at equilibrium, however, when the affinities vanish. As a first approximation, therefore, it is possible to write some linear phenomenological relations between fluxes and affinities:

$$J_\alpha = \sum_\beta L_{\alpha\beta} A_\beta, \tag{9.203}$$

where we have introduced the matrix of the kinetic coefficients $\mathsf{L} = (L_{\alpha\beta})$, whose elements are given by

$$L_{\alpha\beta} = \left. \frac{\partial J_\alpha}{\partial A_\beta} \right|_{F=0}. \tag{9.204}$$

These will generally be a function of the parameters that identify the equilibrium state of the system, such as temperature, pressure, and so on.

We will now see that the matrix L thus introduced can be identified with the matrix L defined earlier. We have, indeed, by substituting eq. (9.200) in eq. (9.203),

$$\frac{\mathrm{d}X_\alpha}{\mathrm{d}t} = \sum_\beta L_{\alpha\beta} A_\beta. \tag{9.205}$$

Multiplying by $X_\gamma(0)$ and taking the correlation at equilibrium, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\langle X_\alpha(t) X_\gamma(0) \right\rangle_{\mathrm{c}} = \sum_\ell L_{\alpha\beta} \left\langle A_\beta(t) X_\gamma(0) \right\rangle_{\mathrm{c}}. \tag{9.206}$$

Here, we have denoted the correlation by $\langle \ldots \rangle_{\mathrm{c}}$, where

$$\langle AB \rangle_{\mathrm{c}} = \langle AB \rangle - \langle A \rangle \langle B \rangle. \tag{9.207}$$

On the other hand, substituting eq. (9.189) we obtain

$$-\sum_{\beta} M_{\alpha\beta} \left\langle X_\beta(t) X_\gamma(0) \right\rangle_{\mathrm{c}} = \sum_{\beta} L_{\alpha\beta} \left\langle A_\beta(t) X_\gamma(0) \right\rangle_{\mathrm{c}} . \tag{9.208}$$

Setting $t = 0$ in this equation, we obtain

$$-\sum_{\beta} M_{\alpha\beta} \left\langle X_\beta X_\gamma \right\rangle_{\mathrm{c}} = \sum_{\beta} L_{\alpha\beta} \left\langle A_\beta X_\gamma \right\rangle_{\mathrm{c}} , \tag{9.209}$$

where the $\left\langle X_\beta X_\gamma \right\rangle_{\mathrm{c}}$, etc., are the instantaneous correlations at equilibrium.

Let us now evaluate the correlations $\left\langle A_\beta X_\gamma \right\rangle_{\mathrm{c}}$. We have

$$\left\langle A_\beta X_\gamma \right\rangle_{\mathrm{c}} = \left\langle \frac{\partial S}{\partial X_\beta} X_\gamma \right\rangle_{\mathrm{c}} - \left\langle \frac{\partial S^{(\mathrm{r})}}{\partial X_\beta^{(\mathrm{r})}} X_\gamma^{(\mathrm{r})} \right\rangle_{\mathrm{c}} . \tag{9.210}$$

Since the reservoir is very large, we can neglect the fluctuations of $\partial S^{(\mathrm{r})} / \partial X_\beta^{(\mathrm{r})}$. Thus the second term vanishes. We remain with

$$\left\langle \frac{\partial S}{\partial X_\beta} X_\gamma \right\rangle_{\mathrm{c}} = \mathcal{N} \int \prod_{\gamma} \mathrm{d}X_\gamma \, \frac{\partial S}{\partial X_\beta} \, X_\gamma \, \mathrm{e}^{S(X_1, \ldots, X_{\mathrm{r}})/k_{\mathrm{B}}} - \left\langle \frac{\partial S}{\partial X_\beta} \right\rangle \left\langle X_\gamma \right\rangle , \tag{9.211}$$

where we have used the fundamental postulate, and introduced the normalization constant $\mathcal{N}$. It is easy to see that $\left\langle \partial S / \partial X_\beta \right\rangle$ vanishes, since the average is evaluated at equilibrium. Integrating the first term by parts, we obtain

$$\left\langle A_\beta X_\gamma \right\rangle = -k_{\mathrm{B}} \, \delta_{\beta\gamma} . \tag{9.212}$$

By substituting, we finally obtain

$$L_{\alpha\gamma} = \frac{1}{k_{\mathrm{B}}} \sum_{\beta} M_{\alpha\beta} \left\langle X_\beta X_\gamma \right\rangle_{\mathrm{c}} . \tag{9.213}$$

Thus, Onsager's reciprocity relations express the symmetry of the matrix L, which relates the affinities to the fluxes.

9.15 Nonequilibrium Steady States

Let us now consider a system that is kept out of equilibrium, by being in contact with two reservoirs, which we denote by (1) and (2). The entropy production in this case is given by

$$\begin{aligned} \dot{S} &= \frac{\mathrm{d}}{\mathrm{d}t} \left(S + S^{(1)} + S^{(2)} \right) \\ &= \sum_{\alpha} \left(\frac{\partial S}{\partial X_\alpha} \frac{\mathrm{d}X_\alpha}{\mathrm{d}t} + \frac{\partial S^{(1)}}{\partial X_\alpha^{(1)}} \frac{\mathrm{d}X_\alpha^{(1)}}{\mathrm{d}t} + \frac{\partial S^{(2)}}{\partial X_\alpha^{(2)}} \frac{\mathrm{d}X_\alpha^{(2)}}{\mathrm{d}t} \right). \end{aligned} \tag{9.214}$$

We assume that the reservoirs are much larger than the system. In this case, most of the entropy is produced in the reservoirs. We can then neglect the contribution of the system

and we obtain

$$\dot{S} \simeq \sum_{\alpha} \left(\frac{\partial S^{(1)}}{\partial X_\alpha^{(1)}} \frac{\mathrm{d}X_\alpha^{(1)}}{\mathrm{d}t} + \frac{\partial S^{(2)}}{\partial X_\alpha^{(2)}} \frac{\mathrm{d}X_\alpha^{(2)}}{\mathrm{d}t} \right) = \sum_{\alpha} \left(\frac{\partial S^{(2)}}{\partial X_\alpha^{(2)}} - \frac{\partial S^{(1)}}{\partial X_\alpha^{(1)}} \right) \frac{\mathrm{d}X_\alpha^{(2)}}{\mathrm{d}t}, \tag{9.215}$$

where we have exploited the fact that, if $\mathrm{d}X_\alpha/\mathrm{d}t$ is negligible, one has

$$\frac{\mathrm{d}X_\alpha^{(1)}}{\mathrm{d}t} = -\frac{\mathrm{d}X_\alpha^{(2)}}{\mathrm{d}t}. \tag{9.216}$$

In this situation, it is natural to define the flux $\mathcal{J}_\alpha$ that goes through the system by

$$\mathcal{J}_\alpha = -\frac{\mathrm{d}X_\alpha^{(1)}}{\mathrm{d}t} = \frac{\mathrm{d}X_\alpha^{(2)}}{\mathrm{d}t}. \tag{9.217}$$

The corresponding affinity is given by

$$\mathcal{A}_\alpha = \frac{\partial S^{(2)}}{\partial X_\alpha^{(2)}} - \frac{\partial S^{(1)}}{\partial X_\alpha^{(1)}}, \tag{9.218}$$

and therefore by the *difference* between the affinities with respect to the two sources. The entropy production therefore assumes the usual form:

$$\dot{S} = \sum_{\alpha} \mathcal{A}_\alpha \mathcal{J}_\alpha. \tag{9.219}$$

We now prove that the entropy production *decreases* with time when the fluxes depend linearly on the affinites. We assume therefore the relation

$$\mathcal{J}_\alpha = \sum_{\beta} L_{\alpha\beta} \mathcal{A}_\beta, \tag{9.220}$$

where $\mathsf{L} = (L_{\alpha\beta})$ is a symmetric, positive definite matrix. This holds if a similar relation holds for each reservoir separately, which requires the fluxes to be small. We also assume that this matrix is constant, at least for the state changes that we want to describe. We then have

$$\begin{aligned} \frac{\mathrm{d}\dot{S}}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t} \sum_{\alpha} \mathcal{A}_\alpha \mathcal{J}_\alpha = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{\alpha\beta} L_{\alpha\beta} \mathcal{A}_\alpha \mathcal{A}_\beta = 2 \sum_{\alpha\beta} L_{\alpha\beta} \frac{\mathrm{d}\mathcal{A}_\alpha}{\mathrm{d}t} \mathcal{A}_\beta \\ &= 2 \sum_{k\ell m} L_{\alpha\beta} \left(\frac{\partial \mathcal{A}_\alpha}{\partial X_\gamma^{(2)}} - \frac{\partial \mathcal{A}_\alpha}{\partial X_\gamma^{(1)}} \right) \mathcal{J}_\gamma \mathcal{A}_\beta \\ &= 2 \sum_{km} \left(\frac{\partial \mathcal{A}_\alpha}{\partial X_\gamma^{(2)}} - \frac{\partial \mathcal{A}_\alpha}{\partial X_\gamma^{(1)}} \right) \mathcal{J}_\alpha \mathcal{J}_\gamma, \end{aligned} \tag{9.221}$$

where we have exploited eq. (9.217) and the relation (9.220). We now observe that

$$\frac{\partial \mathcal{A}_\alpha}{\partial X_\alpha^{(2)}} - \frac{\partial \mathcal{A}_\alpha}{\partial X_\gamma^{(1)}} = \frac{\partial^2 S^{(2)}}{\partial X_\gamma^{(2)} \partial X_\alpha^{(2)}} + \frac{\partial^2 S^{(1)}}{\partial X_\gamma^{(1)} \partial X_\alpha^{(1)}}. \tag{9.222}$$

Due to the concavity of entropy, the matrix $(\partial^2 S/\partial X_\gamma \partial X_\alpha)$ is negative semidefinite for each reservoir. Therefore the expression appearing in the last line of eq. (9.221) cannot

be positive. We thus arrive at the result that the production of entropy $\dot{S}$ decreases as the system evolves. In the present case the system cannot reach equilibrium (at least as long as some of the affinities $\mathcal{A}_\alpha$ are maintained different from zero), but it instead reaches a stationary state, characterized by nonvanishing values of the currents $\mathcal{J}_\alpha$.

We have thus obtained the following variational principle, known as the **principle of minimal entropy production** [Prig47, Prig61]:

> *The nonequilibrium stationary state reached by the system in a linear regime corresponds to the minimum value of the entropy production $\dot{S}$ compatible with the constraints imposed on the system.*

See the work by E. T. Jaynes [Jayn80] for a thorough discussion.

9.16 An Application

In this section, we show a simple application of Onsager's reciprocity relations. Let us consider two containers, containing the same kind of gas, which are in contact by means of a wall that allows for the very slow exchange of gas from one container to the other by means of a throttle (or simply a small hole). The exchange of matter obviously also involves an exchange of internal energy. By focusing our attention on the first container (which we will consider the system), we will be able to write the expressions of the affinities A_E, associated with the internal energy, and A_N, associated with the number of particles:

$$A_E = \frac{\partial S}{\partial E} - \frac{\partial S^{(\mathrm{r})}}{\partial E^{(\mathrm{r})}} = \frac{1}{T} - \frac{1}{T^{(\mathrm{r})}}; \qquad A_N = \frac{\partial S}{\partial N} - \frac{\partial S^{(\mathrm{r})}}{\partial N^{(\mathrm{r})}} = -\left(\frac{\mu}{T} - \frac{\mu^{(\mathrm{r})}}{T^{(\mathrm{r})}}\right). \tag{9.223}$$

The corresponding fluxes are given by

$$J_E = \frac{\mathrm{d}E}{\mathrm{d}t}; \qquad J_N = \frac{\mathrm{d}N}{\mathrm{d}t}. \tag{9.224}$$

We can write the phenomenological linear relations between affinities and fluxes in the form

$$\begin{aligned} J_E &= L_{EE}A_E + L_{EN}A_N; \\ J_N &= L_{NE}A_E + L_{NN}A_N. \end{aligned} \tag{9.225}$$

Thermodynamics and Onsager's relations imply that

$$\begin{aligned} &L_{EE} > 0, \qquad L_{NN} > 0, \qquad L_{EE}L_{NN} - L_{EN}L_{NE} > 0; \\ &L_{EN} = L_{NE}. \end{aligned} \tag{9.226}$$

Let us now consider two situations:

Mechanocaloric effect: The two containers are at the same temperature (but not at the same pressure), and therefore $A_E = 0$. The heat flow between the two containers is given by

$$J_E = L_{EN}A_N = -L_{EN}\Delta\left(\frac{\mu}{T}\right), \tag{9.227}$$

where we have set

$$\Delta\left(\frac{\mu}{T}\right)=\frac{\mu}{T}-\frac{\mu^{(\mathrm{r})}}{T^{(\mathrm{r})}}. \tag{9.228}$$

The particle flow is given by

$$J_N=L_{NN}A_N. \tag{9.229}$$

The ratio of the two flows is given by

$$\frac{J_E}{J_N}=\frac{L_{EN}}{L_{NN}}. \tag{9.230}$$

Thermomechanical effect: The flow of matter between the two containers vanishes: $J_N=0$. This implies the existence of a relation between the affinities in temperature and in the chemical potential:

$$L_{NE}A_E=L_{NE}\,\Delta\left(\frac{1}{T}\right)=-L_{NN}A_N=L_{NN}\,\Delta\left(\frac{\mu}{T}\right), \tag{9.231}$$

where $\Delta(1/T)$ is defined in analogy with $\Delta(\mu/T)$. The ratio between the two affinities is therefore given by

$$-\frac{A_N}{A_E}=\Delta\left(\frac{\mu}{T}\right)\Big/\Delta\left(\frac{1}{T}\right)=\frac{L_{NE}}{L_{NN}}=\frac{L_{EN}}{L_{NN}}, \tag{9.232}$$

where the last relation follows from the Onsager reciprocity relations.

The ratio between the flows in the first case is equal (up to the sign) to the ratio between the affinities in the second. We therefore see how the Onsager reciprocity relations can put in relation two apparently very different phenomena.

In order to make this relation more explicit, let us express the affinity A_N in terms of the difference in pressure. As a result of the Gibbs-Duhem equation, one has

$$\mathrm{d}\mu=-s\,\mathrm{d}T+\frac{1}{\rho}\,\mathrm{d}p, \tag{9.233}$$

where s is the entropy per particle and ρ is the numerical density of the gas. In the case of not too large differences in temperature and pressure, we therefore have

$$A_E=\Delta\left(\frac{1}{T}\right)\simeq-\frac{\Delta T}{T^2}, \tag{9.234}$$

$$A_N=-\Delta\left(\frac{\mu}{T}\right)\simeq-\frac{\Delta\mu}{T}+\mu\,\frac{\Delta T}{T^2}=-\frac{h}{T^2}\,\Delta T+\frac{1}{\rho T}\,\Delta p, \tag{9.235}$$

where $h=\mu+Ts$ is the enthalpy per particle. Thus the ratio between the temperature and the pressure difference in the mechanocaloric effect satisfies the relation

$$\frac{\Delta p}{\Delta T}=\left(h-\frac{L_{EN}}{L_{NN}}\right)\frac{\rho}{T}. \tag{9.236}$$

To evaluate L_{EN}/L_{NN}, let us consider the case in which the contact between the two containers occurs by means of a little hole, much smaller than the mean free path of the gas. In this case, a particle passes from one container to the next conserving the momentum it had in the first container, since there are no impacts that allow it to reequilibrate while passing. If the hole section is S, the number of particles $\mathrm{d}N$ driven by velocity v that exit the first container through the hole in one time unit is given by

$$\mathrm{d}N(\boldsymbol{v}) = \rho S \phi(\boldsymbol{v}) \theta_{\mathrm{H}}(v_x) v_x \, \mathrm{d}\boldsymbol{v}, \tag{9.237}$$

where we have assumed that the hole is parallel to the (yz) plane and that the particles, in order to exit, must move in the direction of positive xs. In this equation, $\phi(\boldsymbol{v})$ is Maxwell's distribution:

$$\phi(\boldsymbol{v}) = \left(\frac{m}{2\pi\, k_{\mathrm{B}} T}\right)^{3/2} \exp\left(-\frac{mv^2}{2k_{\mathrm{B}} T}\right). \tag{9.238}$$

Let us fix the modulus v of the velocity and integrate over its direction. We obtain the number $\mathrm{d}N(v)$ of particles driven by speed v:

$$\mathrm{d}N(v) = \rho S \pi \left(\frac{m}{2\pi k_{\mathrm{B}} T}\right)^{3/2} v^3 \exp\left(\frac{mv^2}{2k_{\mathrm{B}} T}\right) \mathrm{d}v. \tag{9.239}$$

The average energy of a particle that reaches the hole is given by

$$\left\langle \frac{1}{2} mv^2 \right\rangle = \left(\int \mathrm{d}N(v) \, \frac{1}{2} mv^2\right) \Big/ \left(\int \mathrm{d}N(v)\right) = 2k_{\mathrm{B}} T. \tag{9.240}$$

The average transferred energy per particle (in the positive x direction) is therefore $2k_{\mathrm{B}} T$. If the temperatures of the two containers do not differ by much, the ratio between the energy flux and the particle flux is equal to $2k_{\mathrm{B}} T$. We thus obtain

$$\frac{L_{EN}}{L_{NN}} = 2k_{\mathrm{B}} T. \tag{9.241}$$

Recommended Reading

This chapter owes much to chapter 8 of D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford: Oxford University Press, 1987. There is an accessible monograph on the theory of Brownian motion, due to R. M. Mazo, *Brownian Motion: Fluctuations, Dynamics and Applications*, Oxford: Oxford University Press, 2002. The theory of linear response and its connection with correlation functions are very well presented in D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions*, Reading, MA: Benjamin, 1975. One can also usefully read L. E. Reichl, *A Modern Course in Statistical Physics*, Austin: University of Texas Press, 1980; R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II*, Berlin: Springer, 1985; and, for a different point of view, R. Zwanzig, *Nonequilibrium Statistical Mechanics*, Oxford: Oxford University Press, 2001. Stochastic processes such as the random walk have been applied to model financial markets, especially since by Black and

Scholes [Blac73] used it to obtain the expression of the price of an option in a fluctuating market. This has led a growing number of statistical physicists to become interested in the analysis of economic systems, especially in risk management. A useful exposition of the mathematical methods connected with stochastic processes, with its extension to models of financial markets, is due to C. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, 4th ed., Berlin: Springer, 2009. An accessible introduction to the growing field of econophysics is due to R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge, UK: Cambridge University Press, 1999.

I invite everyone to read Einstein's original articles on the theory of Brownian motion, reprinted in A. Einstein, *Investigations on the Theory of Brownian Movement,* New York: Dover, 1956.

10 Stochastic Thermodynamics

> We now see under our eyes heat changed into motion.
> This is contrary to the Second Law of Thermodynamics.
> —attributed to Henri Poincaré

The Fluctuation-Dissipation theorem and Onsager's reciprocal relations allow us to characterize the dynamical behavior of thermodynamic systems close to equilibrium. More recently some relations of more general validity, which hold arbitrarily far from equilibrium, have been discovered. These relations are based on the interpretation of thermodynamic quantities—such as the work performed on a system, the heat exchanged with a reservoir, and the change in a systems entropy—at the level of the single trajectories followed by the system in phase space. This is made possible by a general postulate of **thermodynamic compatibility** and by the separation of the time scales that allows us to describe the dynamics of the system at an intermediate level between the microscopic and the macroscopic, a level known as **mesoscopic**. As a bonus, this approach has led to a better understanding of the thermodynamical implications of measurement and information handling. The present chapter aims at providing a concise introduction to this field of research, very active at present, which has been named **stochastic thermodynamics**.

10.1 Fundamentals

Stochastic thermodynamics applies to **mesoscopic systems** evolving in nonequilibrium in contact with heat (and chemical) reservoirs. Mesoscopic systems are characterized by sizes that, although being small from a macroscopic point of view, are still much larger than atomic sizes. A typical example of such a system is the Brownian particle discussed in sec. 9.1. Other examples are macromolecules, nanodevices, or chemical reactions at very small concentrations. It is required, however, that the dynamics of the mesoscopic system is much slower than the underlying dynamics of the reservoir. Indeed, one often makes the assumption that the heat reservoir relaxes so fast that it can always be considered at equilibrium. This assumption of **timescale separation** is crucial for stochastic thermodynamics to hold. The mesoscopic system can be driven out of equilibrium by **manipulation**,

that is, by a prescribed change of its parameters, or by **driving**, that is, by preventing the approach to equilibrium, for example by attaching the system to more than one reservoir. Both manipulation and driving can be present in a given system.

We assume that the system can be in several states denoted by x, whose energies ϵ_x differ by quantities of order $k_B T$, where T is the temperature of the reservoir. Since the system interacts with the reservoir in unpredictable ways, it follows a stochastic dynamics. The motion of a Brownian particle, described by a Langevin equation, is an example of stochastic dynamics. In the following, unless otherwise specified, we assume that the states x of the system are discrete. The dynamics of a system with discrete states evolving in continuous time is described by a **discrete Markov process**. In such a process, the behavior of the system is dictated (in a statistical sense) by the state it is in, and not by its previous history. More specifically, if the system is in state x' at time t, it can jump to a different state x in a short time interval of duration $\mathrm{d}t$ with a probability $k_{xx'}\,\mathrm{d}t$, where $k_{xx'}$ is called the **jump** (or **transition**) **rate** and depends only on the state x' the system is in at a given time t, on the target state x and possibly on the time t. Then the probability $p_x(t)$ that the system is in state x at time t evolves according the **master equation**[1]

$$\frac{\mathrm{d}p_x}{\mathrm{d}t} = \sum_{x'\,(\neq x)} [k_{xx'} p_{x'}(t) - k_{x'x} p_x(t)]. \tag{10.1}$$

The first term inside brackets describes the rate at which the system performs jumps from other states, landing at x, and the second term describes the rate at which it escapes from x. This suggests to introduce the **escape rate** from each state x by

$$k_x^{\mathrm{out}} = \sum_{x'\,(\neq x)} k_{x'x}. \tag{10.2}$$

One sometimes defines the diagonal elements of the matrix $\mathsf{k} = (k_{xx'})$ by

$$k_{xx} = -k_x^{\mathrm{out}}, \qquad \forall x. \tag{10.3}$$

Then the master equation assumes the simple expression

$$\frac{\mathrm{d}p_x(t)}{\mathrm{d}t} = \sum_{x'} k_{xx'} p_{x'}(t). \tag{10.4}$$

However, we shall not use this notation in the following. With this dynamics, a **trajectory** $\boldsymbol{x}$ of the system is described as a sequence of time intervals in which the system remains in a given state x_k, called **dwells**, separated by **jumps** at discrete times t_k, where the system transitions from, say, x_{k-1} to x_k. We can represent such a trajectory in a time interval $[t_0, t_{\mathrm{f}}]$ by

$$\boldsymbol{x} = (x(t)) = ((x_0, t_0), (x_1, t_1), \ldots, (x_n, t_n), t_{\mathrm{f}}), \tag{10.5}$$

[1] The master equation (G.3) discussed in Appendix G differs from the present one, since it describes evolution in discrete time.

where $x(t) = x_k$ for $t_k \le t < t_{k+1}$, with $t_{n+1} = t_f$. We shall always assume that, given any two different states x and x', there is a finite sequence $(x_0, \ldots, x_{n+1})$ of states such that $x_0 = x$, $x_{n+1} = x'$ and the jump rates $k_{x_{k+1}x_k}$ are positive for $k \in \{0, \ldots, n\}$. We also assume, until further notice, that if $k_{xx'} > 0$, one has $k_{x'x} > 0$, meaning that if a jump $x' \longrightarrow x$ is possible, its reverse $x \longrightarrow x'$ is also possible. With these hypotheses, given the jump rates $\mathsf{K} = (k_{xx'})$, there always is a unique probability distribution $p^{\mathrm{st}} = (p^{\mathrm{st}}_x)$ such that $p^{\mathrm{st}}_x > 0$, $\forall x$, and

$$\sum_{x' \,(\neq x)} \left[k_{xx'}\, p^{\mathrm{st}}_{x'} - k_{x'x}\, p^{\mathrm{st}}_x \right] = 0, \qquad \forall x. \tag{10.6}$$

This can be obtained by applying the methods described in Appendix G to the present case. If the jump rates are time-independent, the probability distribution $p(t) = (p_x(t))$ of the system eventually approaches the distribution p^{st}, that satisfies eq. 10.6. Therefore $p^{\mathrm{st}} = (p^{\mathrm{st}}_x)$ is called the **stationary** or **steady-state** distribution. A special case obtains when the expression in brackets above vanishes for each pair (x, x') of states. In this case there are no net probability currents from one state to another in the steady state. This corresponds to **equilibrium** and the corresponding steady-state probability distribution is denoted by $p^{\mathrm{eq}} = (p^{\mathrm{eq}}_x)$. Then the jump rates satisfy the **detailed balance** condition

$$k_{xx'}\, p^{\mathrm{eq}}_{x'} = k_{x'x}\, p^{\mathrm{eq}}_x, \qquad \forall x, x'. \tag{10.7}$$

Given the jump rates $\mathsf{K}(t) = (k_{xx'}(t))$ (in general time dependent), we can evaluate the probability $\mathcal{P}[\boldsymbol{x}]\, \mathrm{d}t_n \cdots \mathrm{d}t_1$ that the system trajectory is close to a given trajectory $\boldsymbol{x}$ (eq. (10.5)), that is, that the trajectory $((x_0, t_0), (x_1, t'_1), \ldots, (x_n, t'_n), t_f)$ satisfies $0 \le (t'_1 - t_1) \le \mathrm{d}t_1, \ldots, 0 \le (t'_n - t_n) \le \mathrm{d}t_n$. Let us first consider a dwell in state x_k for $t_k \le t < t_{k+1}$. The probability $p^{\mathrm{dwell}}_{x_k}(t_k, t)$ of *not* jumping from x_k before the time t, starting from x_k at t_k, satisfies the differential equation

$$\frac{\mathrm{d}p^{\mathrm{dwell}}_{x_k}}{\mathrm{d}t} = -k^{\mathrm{out}}_{x_k}(t)\, p^{\mathrm{dwell}}_{x_k}, \tag{10.8}$$

with the initial condition $p^{\mathrm{dwell}}_{x_k}(t_k, t_k) = 1$. We obtain therefore

$$p^{\mathrm{dwell}}_{x_k}(t_k, t_{k+1}) = \mathrm{e}^{-\int_{t_k}^{t_{k+1}} \mathrm{d}t\, k^{\mathrm{out}}_{x_k}(t)}. \tag{10.9}$$

On the other hand, the probability that the jump from x_{k-1} to x_k occurs in a time interval of duration $\mathrm{d}t_k$ starting at t_k is given by $k_{x_k x_{k-1}}(t_k)\, \mathrm{d}t_k$. The total probability is obtained by multiplying the probability $p_{x_0}(t_0)$ of the initial condition by the successive factors corresponding to the dwells and the jumps that separate them. It turns out that it is convenient to write these contributions *from right to left*, starting from t_0, followed by the first dwell (t_0, t_1), the first jump at t_1 and so on, to finish with the last dwell from t_n to t_f. We thus obtain

$$\begin{aligned} \mathcal{P}[\boldsymbol{x}]\, \mathrm{d}t_n \cdots \mathrm{d}t_1 = {} & \mathrm{e}^{-\int_{t_n}^{t_f} \mathrm{d}t\, k^{\mathrm{out}}_{x_n}(t)} k_{x_n x_{n-1}}(t_n)\, \mathrm{d}t_n\, \mathrm{e}^{-\int_{t_{n-1}}^{t_n} \mathrm{d}t\, k^{\mathrm{out}}_{x_{n-1}}(t)} \cdots \\ & \times \mathrm{e}^{-\int_{t_1}^{t_2} \mathrm{d}t\, k^{\mathrm{out}}_{x_1}(t)}\, k_{x_1 x_0}(t_1)\, \mathrm{d}t_1\, \mathrm{e}^{-\int_{t_0}^{t_1} \mathrm{d}t\, k^{\mathrm{out}}_{x_0}(t)} p_{x_0}(t_0). \end{aligned} \tag{10.10}$$

The advantage of this layout is that the resulting expression is reminiscent of an element of a matrix product. We have indeed the following expression for the conditional probability $P(x,t;x_0,t_0)$ that the system is in state x at time t, given that it was in state x_0 at time t_0:

$$P(x,t;x_0,t_0)=\sum_{n=0}^{\infty}\sum_{\{x_1,\dots,x_n\}}\left[\delta_{x_n x}\,\mathrm{e}^{-\int_{t_n}^{t_{\mathrm{f}}}\mathrm{d}t\,k^{\mathrm{out}}_{x_n}(t)}\right.\\ \left.\times\prod_{\ell=1}^{n}\left(\int_{t_{\ell-1}}^{t_{\ell+1}}\mathrm{d}t_\ell\;k_{x_\ell x_{\ell-1}}(t_\ell)\,\mathrm{e}^{-\int_{t_{\ell-1}}^{t_\ell}\mathrm{d}t\,k^{\mathrm{out}}_{x_\ell}(t)}\right)\right]. \tag{10.11}$$

10.2 Thermodynamic Consistency and Stochastic Energetics

Let us consider the special case where the jump rates $k_{xx'}$ depend on a parameter λ and, for each value of this parameter, satisfy the detailed balance condition (10.7) with p^{eq}_x equal to the Maxwell-Boltzmann distribution $p^{\mathrm{eq}}_x\propto \mathrm{e}^{-\epsilon_x(\lambda)/k_{\mathrm{B}}T}$. This implies on the one hand that, for each pair such that the jump $x'\longrightarrow x$ is allowed, the reverse jump $x\longrightarrow x'$ is also allowed, and on the other hand that

$$\frac{k_{xx'}(\lambda)}{k_{x'x}(\lambda)}=\mathrm{e}^{(\epsilon_{x'}(\lambda)-\epsilon_x(\lambda))/k_{\mathrm{B}}T}, \tag{10.12}$$

for each pair (x,x') where $k_{xx'}(\lambda)>0$. The difference $\mathcal{Q}_{xx'}=(\epsilon_{x'}(\lambda)-\epsilon_x(\lambda))$ is equal to the amount of energy released to the reservoir in the jump $x'\longrightarrow x$. Since the reservoir is always at equilibrium, this is also equal to $T\,\Delta\mathcal{S}^{\mathrm{res}}_{xx'}$, where T is the reservoir temperature and $\Delta\mathcal{S}^{\mathrm{res}}_{xx'}$ is the change in its entropy due to the jump. We have thus found a connection, in this special case, between the jump rates and the heat released by the system in a jump.

Let us now suppose that the parameter λ is manipulated by the experimenter in a prescribed way: $\boldsymbol{\lambda}=(\lambda(t))$, for $t\in[t_0,t_{\mathrm{f}}]$. We shall refer to $\boldsymbol{\lambda}$ as the **protocol**. If the system dwells in state x for $t_k\le t<t_{k+1}$, its energy changes according to the rule

$$\Delta\epsilon_x=\epsilon_x\big(\lambda(t_{k+1})\big)-\epsilon_x(\lambda(t_k))=\int_{t_k}^{t_{k+1}}\mathrm{d}t\,\frac{\mathrm{d}\lambda(t)}{\mathrm{d}t}\,\frac{\partial\epsilon_x(\lambda(t))}{\partial\lambda}. \tag{10.13}$$

This corresponds to the energy provided to the system by the manipulation during the dwell. We can thus consider it a contribution to the work performed on the system during its evolution. Considering now the whole trajectory $\boldsymbol{x}$, we obtain the relation

$$\Delta\epsilon=\epsilon_{x_{\mathrm{f}}}\big(\lambda(t_{\mathrm{f}})\big)-\epsilon_{x_0}\big(\lambda(t_0)\big)\\ =\underbrace{\sum_{k=0}^{n}\int_{t_k}^{t_{k+1}}\mathrm{d}t\,\frac{\mathrm{d}\lambda(t)}{\mathrm{d}t}\,\frac{\partial\epsilon_x(\lambda(t))}{\partial\lambda}}_{\text{work}}-\underbrace{\sum_{k=1}^{n}\left[\epsilon_{x'}(\lambda(t_k))-\epsilon_x(\lambda(t_k))\right]}_{\text{heat}}. \tag{10.14}$$

Denoting by $\mathcal{W}[\boldsymbol{x}]$ the **stochastic work** performed *on* the system during the evolution along the trajectory $\boldsymbol{x}$, and by $\mathcal{Q}[\boldsymbol{x}]$ the **stochastic heat** released *by* the system on the same trajectory, this relation may be written

$$\Delta\epsilon=\mathcal{W}[\boldsymbol{x}]-\mathcal{Q}[\boldsymbol{x}]. \tag{10.15}$$

This relation is reminiscent of the first principle of thermodynamics. The definitions of work and heat are the opposite of those commonly used in thermodynamics, but they are the most convenient in stochastic thermodynamics, where the stress lies in dissipation and entropy production, rather than in the performance of useful work.

We can represent driving by assuming that the system exchanges energy with an external agent when a jump occurs. In this case, let us denote by $q_{xx'}$ the energy provided by the external agent during the jump $x' \longrightarrow x$. We have $q_{x'x} = -q_{xx'}$. This energy breaks detailed balance and is then immediately released to the reservoir. In this situation, the total energy released to the reservoir in the jump from x' to x is given by

$$\mathcal{Q}_{xx'} = \epsilon_{x'} - \epsilon_x + q_{xx'}. \tag{10.16}$$

Since the heat reservoir cannot distinguish whether the heat it receives comes from a system satisfying detailed balance or not, it is natural to assume that, even in the case of systems with driving, the jump rates satisfy the same condition (10.12), but with $\mathcal{Q}_{xx'}$ given by eq. (10.16):

$$\frac{k_{xx'}(\lambda)}{k_{x'x}(\lambda)} = \mathrm{e}^{(\epsilon_{x'}(\lambda) - \epsilon_x(\lambda) + q_{xx'})/k_{\mathrm{B}}T}. \tag{10.17}$$

This equation represents the **generalized detailed balance** condition and expresses the **thermodynamic consistency** of the dynamics. In the presence of driving, the definitions of stochastic work and stochastic heat must be modified to take into account the driving work $q_{xx'}$:

$$\mathcal{W}[\boldsymbol{x}] = \sum_{k=0}^{n} \int_{t_k}^{t_{k+1}} \mathrm{d}t \, \frac{\mathrm{d}\lambda(t)}{\mathrm{d}t} \, \frac{\partial \epsilon_x(\lambda(t))}{\partial \lambda} + \sum_{k=1}^{n} q_{x_k x_{k-1}}; \tag{10.18}$$

$$\mathcal{Q}[\boldsymbol{x}] = \sum_{k=1}^{n} \mathcal{Q}_{x_k x_{k-1}} = \sum_{k=1}^{n} \left[\epsilon_{x_{k-1}}(\lambda(t_k)) - \epsilon_{x_k}(\lambda(t_k)) + q_{x_k x_{k-1}} \right]. \tag{10.19}$$

We thus recover eq. (10.15). These relations are the core of **stochastic energetics**, in that they provide a consistent interpretation of stochastic work and heat at the level of single trajectories.

10.3 Stochastic Entropy

It is possible to quantify the uncertainty associated with a probability distribution $p = (p_x)$ over the states x. To obtain consistency, one requires that the uncertainty vanish when only one state has a nonvanishing probability; that it be maximal when the distribution is uniform, and that in this case it be monotonically increasing with the number N of states; and that if the system is made of two independent subsystems, it be the sum of the uncertainties relative to each one. It is then possible to show that the only function that satisfies these three requirements (up to a positive multiplicative constant) is the **Shannon**

entropy defined by

$$H_S(p) = -\sum_x p_x \ln p_x. \tag{10.20}$$

The justification of this expression is detailed in Appendix F.1. The Gibbs formula for the entropy (3.136) shows that the entropy of an equilibrium distribution is proportional to its Shannon entropy, where the proportionality constant is the Boltzmann constant k_B:

$$S^{eq} = k_B\, H_S(p^{eq}). \tag{10.21}$$

Since this equation makes sense for an arbitrary probability distribution p, we can exploit it to define the corresponding **nonequilibrium entropy** S by

$$S = k_B\, H_S(p). \tag{10.22}$$

Equation (10.20) can be interpreted as the average over the distribution p of the random variable $\ln p_x$. This opens the way to define the **stochastic entropy** $\mathcal{S}_x$ for a system characterized by a probability distribution p:

$$\mathcal{S}_x = -k_B\, \ln p_x. \tag{10.23}$$

The nonequilibrium entropy S is then obtained as the average of the stochastic entropy $\mathcal{S}_x$:

$$S = \sum_x p_x\, \mathcal{S}_x = \langle \mathcal{S}_x \rangle_p. \tag{10.24}$$

The dissimilarity between two probability distributions, p and q, of the same system is quantified by the **Kullback-Leibler divergence** (also called the **relative entropy**), defined by

$$D_{KL}(p\|q) = \sum_x p_x \ln \frac{p_x}{q_x}. \tag{10.25}$$

As shown in sec. F.1, we have $D_{KL}(p\|q) \geq 0$, and $D_{KL}(p\|q) = 0$ if and only if $p_x = q_x$, $\forall x$. However, one has in general $D_{KL}(p\|q) \neq D_{KL}(q\|p)$. We can take advantage of the Kullback-Leibler divergence to prove the variational principle for the canonical ensemble discussed in sec. 3.17, that is, that the entropy S is maximal for the equilibrium distribution p^{eq}, from among all distributions with the same average value of the energy. Since $p_x^{eq} = e^{-(\epsilon_x - F)/k_B T}$, where $F = E - T\, S^{eq}$ is the Helmholtz free energy, we have

$$\begin{aligned} 0 \leq D_{KL}(p\|p^{eq}) &= \sum_x p_x\, \ln \frac{p_x}{p_x^{eq}} = -H_S(p) + \sum_x p_x\, \frac{\epsilon_x - F}{k_B T} \\ &= -H_S(p) + \frac{\langle \epsilon \rangle_p - F}{k_B T}. \end{aligned} \tag{10.26}$$

Since $\langle \epsilon \rangle_p = E$ by our assumption, multiplying by k_B and using eq. (10.22) we obtain

$$S^{eq} - S \geq 0. \tag{10.27}$$

Let us now consider a mesoscopic system, prepared with a well-defined initial distribution $p(t_0)$ and evolving under manipulation and/or driving in the time interval $[t_0, t_f]$. In this condition, the probability distribution $p(t) = (p_x(t))$ for $t_0 \le t \le t_f$ is obtained as the solution of the master equation (10.1) with the initial condition p_0. If the system is in state x at time t, we can assign to it the stochastic entropy $\mathcal{S}_x(t) = -k_B \ln p_x(t)$.

Given the trajectory $\boldsymbol{x}$ and the manipulation protocol $\boldsymbol{\lambda}$, the entropy change of the system between the initial and the final state is given by

$$\Delta \mathcal{S}^{\rm sys} = \mathcal{S}_{x_f}(t_f) - \mathcal{S}_{x_0}(t_0), \tag{10.28}$$

and the total entropy change of the system plus the heat reservoir is given by

$$\Delta \mathcal{S}^{\rm tot}[\boldsymbol{x}] = \Delta \mathcal{S}^{\rm sys} + \Delta \mathcal{S}^{\rm res}[\boldsymbol{x}], \tag{10.29}$$

where, according to the assumption of fast equilibration of the reservoir, its entropy change is given by

$$\Delta \mathcal{S}^{\rm res}[\boldsymbol{x}] = T\, \mathcal{Q}[\boldsymbol{x}], \tag{10.30}$$

with $\mathcal{Q}[\boldsymbol{x}]$ given by eq. (10.19).

Exercise 10.1 Consider a system in contact with a reservoir at the temperature T, but in a nonequilibrium state described by the probability distribution $p = (p_x)$, when the equilibrium distribution $p^{\rm eq}$ is the canonical one, $p_x^{\rm eq} = \mathrm{e}^{(F-\epsilon_x)/k_B T}$. Show that the work W that can be extracted by allowing the system to relax to equilibrium satisfies the inequality

$$W \le k_B T D_{\rm KL}(p \| p^{\rm eq}),$$

where $D_{\rm KL}(p\|q)$ is the Kullback-Leibler divergence between the probability distributions p and q.

Given a system satisfying the master equation

$$\frac{\mathrm{d}p_x}{\mathrm{d}t} = \sum_{x' (\neq x)} [k_{xx'} p_{x'}(t) - k_{x'x} p_x(t)],$$

where the matrix $\mathsf{K} = (k_{xx'})$ of transition rates is ergodic (see sec. G.3), and therefore has a unique stationary distribution $p^{\rm st} = (p_x^{\rm st})$, one has

$$\frac{\mathrm{d}}{\mathrm{d}t} D_{\rm KL}(p(t) \| p^{\rm st}) \le 0,$$

where $p(t) = (p_x(t))$ is the solution of the master equation. Thus, the system evolves toward the stationary state described by $p^{\rm st}$.

To prove this result, define, for $0 \le x < \infty$,

$$f(x) = x \ln x. \tag{10.31}$$

We then have

$$D_{\mathrm{KL}}(p(t)\|p^{\mathrm{st}}) = \sum_x p_x^{\mathrm{st}} f\left(\frac{p_x(t)}{p_x^{\mathrm{st}}}\right). \tag{10.32}$$

Define now $w_x = p_x/p_x^{\mathrm{st}}$, $\forall x$. We then have

$$\frac{\mathrm{d}}{\mathrm{d}t} D_{\mathrm{KL}}\,(p(t)\|p^{\mathrm{st}}) = \sum_{x\neq x'} k_{xx'} p_{x'}^{\mathrm{st}}\,[w_{x'} f'(w_x) - w_{x'} f'(w_{x'})]. \tag{10.33}$$

We now have, for an arbitrary vector $\psi = (\psi_x)$,

$$\sum_{x\neq x'} k_{xx'} p_{x'}^{\mathrm{st}}\,(\psi_x - \psi_{x'}) = 0. \tag{10.34}$$

Choose

$$\psi_x = f(w_x) - w_x f'(w_x), \qquad \forall x. \tag{10.35}$$

We then have, inserting $(\psi_x - \psi_{x'})$ in each term of eq. (10.33),

$$\frac{\mathrm{d}}{\mathrm{d}t} D_{\mathrm{KL}}(p(t)\|p^{\mathrm{st}}) = \sum_{xx'} k_{xx'} p_{x'}^{\mathrm{st}}\left[f(w_x) + (w_{x'} - w_x) f'(w_x) - f(w'_x)\right]. \tag{10.36}$$

Now, since the function $f(x)$ is convex, the quantity in brackets cannot be positive (the curve $f(x)$ lies always above any of its tangents). We obtain therefore

$$\frac{\mathrm{d}}{\mathrm{d}t} D_{\mathrm{KL}}(p(t)\|p^{\mathrm{st}}) \leq 0, \tag{10.37}$$

and the inequality is strict, since $f(x)$ is strictly convex, if for any pair (x, x') for which $k_{xx'} > 0$ one has $w_x \neq w_{x'}$. Thus, since $D_{\mathrm{KL}}(p\|p^{\mathrm{st}}) \geq 0$, it converges to its minimum, corresponding to $w_x = \mathrm{const.} = 1$ (by normalization), that is, $p_x = p_x^{\mathrm{st}}$, $\forall x$. The case of a non-ergodic transition-rate matrix K is left to the reader.

Exercise 10.2 Prove eq. (10.34).

Exercise 10.3 Discuss the case of a non-ergodic transition rate matrix K.

Exercise 10.4 Find an alternate proof of the inequality (10.37) when $p(t)$ is a solution of a master equation defined by the ergodic matrix $\mathsf{K} = (k_{xx'})$ satisfying detailed balance (that is, $k_{xx'} p_{x'}^{\mathrm{eq}} = k_{x'x} p_x^{\mathrm{eq}}$, $\forall x, x'$), and $p^{\mathrm{eq}} = (p_x^{\mathrm{eq}})$ is its equilibrium solution.

10.4 Simple Examples

A simple example of a manipulated system is a two-level system with $x \in \{0, 1\}$. We denote the jump rates by $k^{\downarrow} = k_{01} = b_0$ and $k^{\uparrow}(t) = k_{10}(t) = a_0\,[1 + \lambda(t)]$, where $\lambda(t) = \gamma\,\sin(2\pi\, t/t_{\mathrm{m}})$. This system has been experimentally realized in a diamond-defect system. The results of

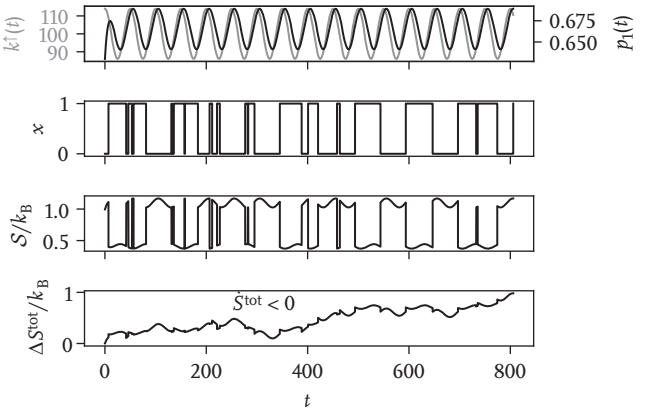


FIGURE 10.1. Simulation of the two-level system with parameters: $a_0^{-1} = 10\,\text{ms}$, $b_0^{-1} = 20\,\text{ms}$, $\gamma = 0.14$, $t_m = 50\,\text{ms}$. Time measured in ms. Top to bottom: Jump rate $k^\uparrow(t)$ in kHz as a function of time (gray) and occupation probability $p_1(t)$ (black); a trajectory of the system; stochastic entropy $\mathcal{S}_{x(t)}(t)$ for the given trajectory, and the corresponding total produced entropy ΔS^{tot}. The entropy production appears to be negative between $t = 250\,\text{ms}$ and $t = 320\,\text{ms}$.

a straightforward simulation are shown in fig. 10.1. It is interesting to notice that in this realization, the curve representing the total entropy produced since the beginning of the experiment exhibits a pronounced dip, corresponding to a negative entropy production rate, between $t = 250\,\text{ms}$ and $t = 320\,\text{ms}$. This can be considered as a fleeting violation of the second law.

A system with three or more states x can exhibit driving. Let us thus consider a three-state system with $x \in \{0, 1, 2\}$. Solving a simple linear system we obtain the steady state as a function of the jump rates $k_{xx'}$:

$$\begin{aligned} p_0^{st} &= \mathcal{N}\,(k_{01}k_{02} + k_{01}k_{12} + k_{02}k_{21})\,; \\ p_1^{st} &= \mathcal{N}\,(k_{10}k_{12} + k_{10}k_{02} + k_{12}k_{20})\,; \\ p_2^{st} &= \mathcal{N}\,(k_{20}k_{21} + k_{20}k_{01} + k_{21}k_{10}). \end{aligned} \tag{10.38}$$

Here $\mathcal{N}$ is a normalization constant:

$$\mathcal{N}^{-1} = k_{01}k_{02} + k_{01}k_{12} + k_{02}k_{21} + k_{10}k_{12} + k_{10}k_{02} + k_{12}k_{20} + k_{20}k_{21} + k_{20}k_{01} + k_{21}k_{10}. \tag{10.39}$$

To be more specific, we assign to the states the energies $\epsilon_0 = 0$, ϵ_1, and ϵ_2, with $\epsilon_1 > \epsilon_2 > 0$. We also assume that there is a driving q that favors the $2 \longrightarrow 1$ transition. We thus have, by the generalized detailed balance condition,

$$\begin{aligned} \frac{k_{01}}{k_{10}} &= \mathrm{e}^{-\epsilon_1/k_B T}; \\ \frac{k_{02}}{k_{20}} &= \mathrm{e}^{-\epsilon_2/k_B T}; \\ \frac{k_{12}}{k_{21}} &= \mathrm{e}^{(\epsilon_2 - \epsilon_1 + q)/k_B T}. \end{aligned} \tag{10.40}$$

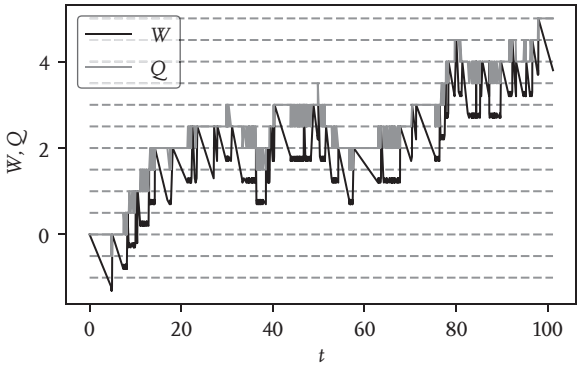


FIGURE 10.2. Stochastic work $\mathcal{W}$ and heat $\mathcal{Q}$ in a simulation of a three-state system with $\epsilon_1 = 1.2\, k_B T$, $\epsilon_2 = 0.8\, k_B T$, $q = 0.5\, k_B T$, $\omega = 1$, and $\omega_1 = 5$ (arbitrary units). The dashed lines are at intervals of q; therefore each full cycle $0 \longrightarrow 2 \longrightarrow 1 \longrightarrow 0$ direction corresponds to going up one step. Reverse motion takes place for t between 50 and 60.

These relations do not full specify the jump rates. One possibility is to set $k_{01} = k_{02} = \omega$, $k_{21} = \omega_1$, with $\omega_1 > \omega$. In this situation we expect the system to spend most of its time hovering between the states 1 and 2, and eventually leaving them more often via the $1 \longrightarrow 0$ pathway than via the $2 \longrightarrow 0$ pathway. Each full cycle $0 \longrightarrow 2 \longrightarrow 1 \longrightarrow 0$ corresponds to a net amount of heat q released to the heat reservoir, that is, in a net entropy production $\Delta S^{\text{tot}} = q/T$. The results of a simulation are shown in fig. 10.2.

10.5 Average Entropy Production Rate

Let us now discuss the average rate of entropy production. During each jump from state x' to state x an amount $\mathcal{Q}_{xx'}$ of heat, given by eq. (10.16), is released to the reservoir. This corresponds to an increment $\Delta\mathcal{S}^{\text{res}}_{xx'}$ of its entropy, given by

$$\Delta\mathcal{S}^{\text{res}}_{xx'} = k_B \ln \frac{k_{xx'}}{k_{x'x}}, \tag{10.41}$$

where we have taken advantage of eq. (10.17). During the same jump, the system entropy changes by

$$\Delta\mathcal{S}^{\text{sys}}_{xx'} = k_B \frac{\ln p_{x'}}{\ln p_x}. \tag{10.42}$$

Therefore, the total entropy change during a jump is given by

$$\Delta\mathcal{S}^{\text{tot}}_{xx'} = \Delta\mathcal{S}^{\text{res}}_{xx'} + \Delta\mathcal{S}^{\text{sys}}_{xx'} = k_B \ln \frac{k_{xx'} p_{x'}}{k_{x'x} p_x}. \tag{10.43}$$

On the other hand, during a dwell in the state x for, say, $t_k \le t < t_{k+1}$, there is no heat release, and therefore the reservoir entropy does not change. The system entropy changes instead, because the probability $p_x(t)$ is in general time-dependent. Therefore the total entropy

change during a dwell is given by

$$\Delta \mathcal{S}^{\mathrm{sys}} = k_{\mathrm{B}} \left(\ln p_x(t_k) - \ln p_x(t_{k+1}) \right) = -k_{\mathrm{B}} \int_{t_k}^{t_{k+1}} \mathrm{d}t \, \frac{1}{p_x} \frac{\mathrm{d}p_x}{\mathrm{d}t}. \tag{10.44}$$

Because of the conservation of probability, this change vanishes on average:

$$\left\langle \frac{1}{p_x} \frac{\mathrm{d}p_x}{\mathrm{d}t} \right\rangle = \sum_x p_x \, \frac{1}{p_x} \frac{\mathrm{d}p_x}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \sum_x p_x = 0. \tag{10.45}$$

Therefore the net average entropy production is due only to the jumps. We can obtain the net probability current from x' to x by multiplying the jump rate $k_{xx'}(t)$ by the probability $p_x(t)$ of the state x' and subtracting the contribution of the opposite jump:

$$J_{xx'}(t) = k_{xx'}(t) p_{x'}(t) - k_{x'x}(t) p_x(t). \tag{10.46}$$

Then the average entropy production rate is obtained by summing, for each state pair (x, x'), the total entropy change $\Delta \mathcal{S}^{\mathrm{tot}}_{xx'}$ multiplied by the net probability current $J_{xx'}$. We obtain

$$\dot{S}(t) = \frac{k_{\mathrm{B}}}{2} \sum_{x \neq x'} J_{xx'}(t) \, \ln \frac{k_{xx'}(t) p_{x'}(t)}{k_{x'x}(t) p_x(t)}, \tag{10.47}$$

where the factor 1/2 compensates for the double counting. Now, since the sign of $\ln(k_{xx'}(t)p_{x'}(t)) - \ln(k_{x'x}(t)p_x(t))$ is the same as that of $k_{xx'}(t)p_{x'}(t) - k_{x'x}(t)p_x(t)$, each term in this sum is nonnegative. Clearly, it can only vanish if the system satisfies the detailed balance condition (10.12) and the probability p is the equilibrium one p^{eq}. Equation (10.47) was obtained by J. Schnakenberg [Schn76], developing ideas introduced by Bergmann and Lebowitz [Berg55], and is known as the **Schnakenberg formula**.

10.6 Time Reversal

Let us now consider a system manipulated according to a protocol $\boldsymbol{\lambda}$ defined for $t_0 \leq t \leq t_{\mathrm{f}}$. To each trajectory $\boldsymbol{x} = (x(t))$ we can associate its time-reversed image $\widehat{\boldsymbol{x}} = (\widehat{x}(t))$ such that $\widehat{x}(t) = x(\hat{t})$, where $\hat{t} = t_0 + t_{\mathrm{f}} - t$.[2] We shall call it the **backward trajectory**. In the same way, we define the **backward protocol** $\widehat{\boldsymbol{\lambda}}$ by $\widehat{\lambda}(t) = \lambda(\hat{t})$, $t_0 \leq t \leq t_{\mathrm{f}}$.

Let us now evaluate the ratio $\mathcal{P}[\boldsymbol{x}; \boldsymbol{\lambda}] / \mathcal{P}[\widehat{\boldsymbol{x}}; \widehat{\boldsymbol{\lambda}}]$ of the probability of the trajectory $\boldsymbol{x}$ (the **forward trajectory**) under the protocol $\boldsymbol{\lambda}$ (the **forward protocol**) to that of the backward trajectory $\widehat{\boldsymbol{x}}$ with the backward protocol $\widehat{\boldsymbol{\lambda}}$. By the definition of the backward protocol, one sees that, for any pair (x, x') of states x, the jump rate $k_{xx'}(\lambda(t))$ satisfies

$$k_{xx'}(\lambda(t)) = k_{xx'}(\widehat{\lambda}(\hat{t})), \qquad \forall x, x', \quad t_0 \leq t \leq t_{\mathrm{f}}. \tag{10.48}$$

[2] In this chapter, we assume that the time-reversed image of a state x is equal to x. This would not be the case, for example, in a Hamiltonian system in which the state is defined by positions and momenta of the particles. See, for example, sec. 9.11.

As a consequence of this relation, one has $k_x^{\text{out}}(\lambda(t)) = k_x^{\text{out}}(\widehat{\lambda}(\hat{t}))$, $\forall x$, and therefore the contribution of the dwells cancel out between the numerator and the denominator. We are then left with

$$\frac{\mathcal{P}[\boldsymbol{x};\boldsymbol{\lambda}]}{\mathcal{P}[\widehat{\boldsymbol{x}};\widehat{\boldsymbol{\lambda}}]} = \prod_{k=1}^{n} \frac{k_{x_k x_{k-1}}(\lambda(t_k))}{k_{x_{k-1}x_k}(\lambda(t_k))} \cdot \frac{p_{x_0}(t_0)}{p^{(\mathrm{B})}_{x_{\mathrm{f}}}(\hat{t}_0)}, \tag{10.49}$$

where we have denoted by $p^{(\mathrm{B})}_{x_{\mathrm{f}}}(\hat{t}_0)$ the initial probability distribution of the backward trajectory. Things simplify if we take

$$p^{(\mathrm{B})}_{x_{\mathrm{f}}}(\hat{t}_0) = p_{x_{\mathrm{f}}}(t_{\mathrm{f}}), \tag{10.50}$$

but it is easy to accommodate more general hypotheses. Taking into account eqs. (10.17), (10.19), and (10.28) we obtain

$$\frac{\mathcal{P}[\boldsymbol{x};\boldsymbol{\lambda}]}{\mathcal{P}[\widehat{\boldsymbol{x}};\widehat{\boldsymbol{\lambda}}]} = \exp\left[\frac{1}{k_{\mathrm{B}}}\left(\sum_{k=1}^{n} \mathcal{Q}_{x_k x_{k-1}} + \Delta\mathcal{S}^{\mathrm{sys}}\right)\right] = \exp\left(\frac{\Delta\mathcal{S}^{\mathrm{tot}}[\boldsymbol{x};\boldsymbol{\lambda}]}{k_{\mathrm{B}}}\right), \tag{10.51}$$

where we have exploited eq. (10.29).

This is one of the fundamental relations of stochastic thermodynamics. It explicitly connects, for each trajectory, the entropy production to the violation of time-reversal invariance.

10.7 Fluctuation Relations

An immediate consequence of eq. (10.51) is obtained by dividing by the right-hand side and multiplying by $\mathcal{P}[\widehat{\boldsymbol{x}};\widehat{\boldsymbol{\lambda}}]$. Since the correspondence between $\boldsymbol{x}$ and $\widehat{\boldsymbol{x}}$ is one-to-one, and the volume elements satisfy $\mathcal{D}\boldsymbol{x} = \mathcal{D}\widehat{\boldsymbol{x}}$, we can integrate over $\boldsymbol{x}$. We have by definition, for any functional $\mathcal{A}[\boldsymbol{x}]$ of the trajectory $\boldsymbol{x}$,

$$\int \mathcal{D}\boldsymbol{x}\, \mathcal{A}[\boldsymbol{x}] = \sum_{n=0}^{\infty} \sum_{\{x_0, x_1, \ldots, x_n\}} \int_{t_0}^{t_{\mathrm{f}}} \mathrm{d}t_1 \int_{t_1}^{t_{\mathrm{f}}} \mathrm{d}t_2 \cdots \int_{t_{n-1}}^{t_{\mathrm{f}}} \mathrm{d}t_n\, \mathcal{A}[\boldsymbol{x}]. \tag{10.52}$$

By the normalization of $\boldsymbol{\mathcal{P}}[\widehat{\boldsymbol{x}};\widehat{\boldsymbol{\lambda}}]$ we obtain

$$\int \mathcal{D}\boldsymbol{x}\, \mathcal{P}[\boldsymbol{x};\boldsymbol{\lambda}]\, \mathrm{e}^{-\Delta\mathcal{S}^{\mathrm{tot}}[\boldsymbol{x};\boldsymbol{\lambda}]/k_{\mathrm{B}}} = \left\langle \mathrm{e}^{-\Delta\mathcal{S}^{\mathrm{tot}}[\boldsymbol{x};\boldsymbol{\lambda}]/k_{\mathrm{B}}} \right\rangle_{\boldsymbol{\lambda}} = 1. \tag{10.53}$$

This equality is called the **integral fluctuation relation**. As a first consequence, it implies that if the system is out of equilibrium, so that the probability of having a positive entropy production is nonvanishing, there also is a positive probability of observing *negative* entropy production. Indeed, since $\mathrm{e}^{-\Delta\mathcal{S}^{\mathrm{tot}}/k_{\mathrm{B}}} < 1$ for $\Delta\mathcal{S}^{\mathrm{tot}} > 0$, the equality could not be satisfied if only positive values of $\Delta\mathcal{S}^{\mathrm{tot}}$ were obtained. On the other hand, for any real-valued random variable x, we have by Jensen's inequality

$$\langle \mathrm{e}^{x} \rangle \geq \mathrm{e}^{\langle x \rangle}. \tag{10.54}$$

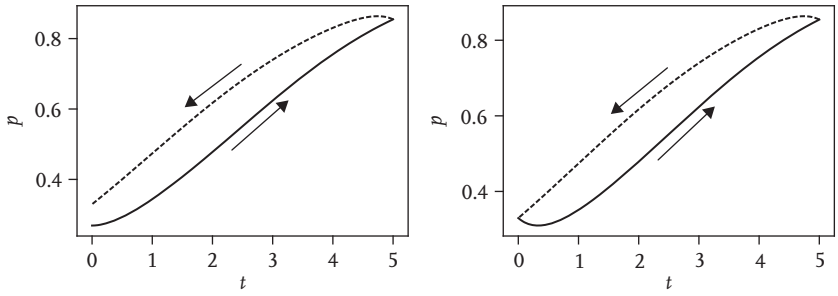


FIGURE 10.3. Probability $p(t)$ of the 0 state for a linear manipulation between $\lambda_0 = -k_B T$ and $\lambda_f = 2\, k_B T$ of duration $t_f = 5/\omega$. Full line: Forward manipulation. Dashed line: Reverse manipulation. Left panel: $p(0) = 1/(1 + e^{-\lambda_0/k_B T})$. Involution does not hold. Right panel: With a suitably chosen $p(0)$, one has $p(t_f; \widehat{\boldsymbol{\lambda}}) = p(0)$, and the involution property holds.

This relation is derived in Appendix B.1. We obtain therefore

$$\langle \Delta \mathcal{S}^{\rm tot} \rangle \geq 0. \tag{10.55}$$

This inequality can be considered as the second law of stochastic thermodynamics. While the first law, expressed by eq. (10.15), holds trajectory by trajectory, this relation shows that the second law holds only on average.

If the initial conditions are suitably chosen, the entropy production $\Delta \mathcal{S}^{\rm tot}[\boldsymbol{x}; \boldsymbol{\lambda}]$ exhibits the **involution property**

$$\Delta \mathcal{S}^{\rm tot}[\boldsymbol{x}; \boldsymbol{\lambda}] = -\Delta \mathcal{S}^{\rm tot}[\widehat{\boldsymbol{x}}; \widehat{\boldsymbol{\lambda}}]. \tag{10.56}$$

For example, in fig. 10.3 we show the evolution of a two-level system with energies $\epsilon_0 = 0$ and $\epsilon_1(t) = \lambda_0 + (\lambda_1 - \lambda_0) t/t_f$, with the jump rates satisfying detailed balance:

$$k_{01}(t) = \omega\, e^{\epsilon_1(t)/2k_B T}; \qquad k_{10}(t) = \omega\, e^{-\epsilon_1(t)/2k_B T}. \tag{10.57}$$

The plot represents the probability p of being in the 0 state, on a manipulation $\boldsymbol{\lambda}$ with $\lambda_0 = -k_B T$, $\lambda_f = 2\, k_B T$, and $\omega\, t_f = 5$. The continuous line is the solution upon a forward manipulation, and the dashed line the inverse trajectory (plotted against $\hat{t} = t_f - t$) with the reverse manipulation, starting from the end state of the forward manipulation. In the left panel, the initial condition is $p(t_0) = 1/(1 + e^{-\lambda_0/k_B T})$, and the final state of the reverse evolution is different from the initial state, and the involution property does not hold. In the right panel, the initial condition $p(t_0)$ is judiciously chosen, so that the state at the end of the manipulation is equal to the initial state of the forward manipulation, and the involution property holds.

If the involution property holds, we can integrate (10.53) over all trajectories $\boldsymbol{x}$ with a given value S of $\Delta \mathcal{S}^{\rm tot}$, obtaining

$$\frac{P(S; \boldsymbol{\lambda})}{P(-S; \widehat{\boldsymbol{\lambda}})} = e^{S/k_B}. \tag{10.58}$$

This is known as the **detailed fluctuation relation**. Equation (10.51) is also considered an expression of a detailed fluctuation relation. One especially important case obtains when

the system is in a nonequilibrium steady state. In this case the involution property holds trivially, and the typical values of $\Delta\mathcal{S}^{\text{tot}}$ increase proportionally to the duration $\mathcal{T} = t_{\text{f}} - t_0$ of the observed time interval. Setting $S = \mathcal{T}\,\dot{S}$, where $\dot{S}$ is the mean entropy production rate in the interval, we obtain

$$\frac{P_{\mathcal{T}}(\dot{S})}{P_{\mathcal{T}}(-\dot{S})} = \mathrm{e}^{\mathcal{T}\,\dot{S}/k_{\text{B}}}, \tag{10.59}$$

where $P_{\mathcal{T}}(\dot{S})$ is the probability density of $\dot{S}$ over a time interval of duration $\mathcal{T}$. Clearly, the probability of observing negative entropy production rates decreases exponentially with $\mathcal{T}$. Equation (10.59) is historically one of the first fluctuation relations to have been found, and is known as the **Gallavotti-Cohen relation** [Gall95]. It was first derived in the context of Hamiltonian dynamics, and generalized to stochastic dynamics by Lebowitz and Spohn [Lebo99] and Kurchan [Kurc98].

Another interesting case is when the system is manipulated in the absence of driving. It is useful to define the manipulation protocol $\boldsymbol{\lambda}$ in the following way: starting from, say, $t = 0$ we keep the value of λ fixed at λ_0 until the time t_0. We then perform a prescribed manipulation until reaching the value λ_{f} at time t_{f}. Later, we keep the value of λ fixed at λ_{f} until the time t_{end}. We assume that t_0 and $t_{\text{end}} - t_{\text{f}}$ are large enough, so that the system is able to relax to equilibrium in either interval. Then the system evolves from an initial equilibrium distribution at $t = 0$ to a final equilibrium distribution at time t_{end}, but this equilibrium distribution is also the initial equilibrium distribution at t_0 of the backward manipulation, and the involution relation is satisfied. Moreover, the change of the system entropy from the beginning to the end of the manipulation is given by

$$\begin{aligned}\Delta\mathcal{S}^{\text{sys}} &= -k_{\text{B}}\left(\ln p^{\text{eq}}_{x_{\text{f}}}(\lambda_{\text{f}}) - \ln p^{\text{eq}}_{x_0}(\lambda_0)\right)\\ &= -\frac{1}{T}\left(\epsilon_{x_0}(\lambda_0) - F(\lambda_0) - \epsilon_{x_{\text{f}}} + F(\lambda_{\text{f}})\right),\end{aligned} \tag{10.60}$$

where $F(\lambda)$ is the Helmholtz free energy corresponding to the value λ of the parameter. Substituting in eq. (10.29) and taking into account eq. (10.15) we obtain

$$T\,\Delta\mathcal{S}^{\text{tot}}[\boldsymbol{x};\boldsymbol{\lambda}] = \mathcal{W}[\boldsymbol{x};\boldsymbol{\lambda}] - \Delta F, \tag{10.61}$$

where $\Delta F = F(\lambda_{\text{f}}) - F(\lambda_0)$. Substituting in the integral fluctuation relation, we obtain the **Jarzynski equality**

$$\langle \mathrm{e}^{-\mathcal{W}/k_{\text{B}}T}\rangle = \mathrm{e}^{-\Delta F/k_{\text{B}}T}. \tag{10.62}$$

By eq. (10.54), this relation implies in particular that the dissipated work $\mathcal{W}^{\text{diss}} = \mathcal{W} - \Delta F$ is positive on average. One has to stress that the contribution to $\mathcal{W}$ comes only from the time interval between t_0 and t_{f}, while the time intervals between 0 and t_0, or between t_{f} and t_{end}, do not contribute. In particular, the only requirement is that the system be in equilibrium at the beginning of both the forward and the backward manipulation, but it has no need to be in equilibrium at t_{f} in either. Therefore this relation allows us to express the difference ΔF of the equilibrium free energies by nonequilibrium experiments. Such

experiments have been carried on, for example, by unfolding and refolding RNA hairpins, manipulating them via optical tweezers. A more reliable evaluation of this quantity can be obtained by the corresponding detailed fluctuation relation, obtained by summing over all trajectories with the same value W of the stochastic work $\mathcal{W}$. We thus obtain the **Crooks relation**

$$\frac{P(W;\boldsymbol{\lambda})}{P(-W;\widehat{\boldsymbol{\lambda}})} = \mathrm{e}^{(W-\Delta F)/k_{\mathrm{B}}T}. \tag{10.63}$$

Clearly, this relation implies that ΔF is the value of W where the probabilities $P(W;\boldsymbol{\lambda})$ and $P(-W;\widehat{\boldsymbol{\lambda}})$ cross.

Exercise 10.5 Show that the value W of the work that contributes the most to the Jarzynski average $\langle \mathrm{e}^{-\mathcal{W}/k_{\mathrm{B}}T}\rangle$ is equal to the opposite of the most probable value of W in the backward protocol $\widehat{\boldsymbol{\lambda}}$.

Exercise 10.6 Show that the Jarzynski equality holds for a system undergoing an instantaneous quench, defined by the protocol

$$\lambda(t) = \begin{cases} \lambda_0, & \text{for } 0 \le t_0 < t_{\mathrm{q}}; \\ \lambda_{\mathrm{f}}, & \text{for } t_{\mathrm{q}} \le t \le t_{\mathrm{f}}. \end{cases}$$

Assume that the system is at equilibrium with $\lambda = \lambda_0$ at the temperature T for $t = t_0$.

Exercise 10.7 Show that, if the entropy production $\Delta\mathcal{S}^{\mathrm{tot}}$ has a distribution, one has

$$2k_{\mathrm{B}}\left\langle \Delta\mathcal{S}^{\mathrm{tot}}\right\rangle = \sigma^2_{\mathrm{tot}},$$

where σ^2_{tot} is the variance of $\Delta\mathcal{S}^{\mathrm{tot}}$.

Exercise 10.8 (Onsager reciprocal relations) In a nonequilibrium steady state produced by the affinities $\mathcal{A} = (\mathcal{A}_\alpha)$, the entropy produced during a time interval of duration $\mathcal{T}$ is given by eq. (9.219),

$$\mathcal{S} = \mathcal{T}\sum_\alpha \mathcal{A}_\alpha \mathcal{J}_\alpha,$$

where $\mathcal{J} = (\mathcal{J}_\alpha)$ are the fluctuating currents conjugate to the affinities $\mathcal{A} = (\mathcal{A}_\alpha)$.

Using the detailed fluctuation relation, show that the quantity

$$\psi(q) = \lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \ln\left\langle \exp\left(\mathcal{T}\sum_\alpha q_\alpha \mathcal{J}_\alpha\right)\right\rangle,$$

where $q = (q_\alpha)$, possesses the Gallavotti-Cohen symmetry $q_\alpha - A_\alpha/k_{\mathrm{B}} \longrightarrow -q_\alpha$, that is,

$$\psi\left(q - \frac{\mathcal{A}}{k_{\mathrm{B}}}\right) = \psi(-q).$$

Using this result and the relations

$$\langle \mathcal{J}_\alpha \rangle = \left. \frac{\partial \psi}{\partial q_\alpha} \right|_{q=0}, \qquad L_{\alpha\beta} = \left. \frac{\partial \langle \mathcal{J}_\alpha \rangle}{\partial \mathcal{A}_\alpha} \right|_{q=0,A=0} = \left. \frac{\partial^2 \psi}{\partial q_\alpha \, \partial \mathcal{A}_\alpha} \right|_{q=0,A=0},$$

prove that the matrix $\mathsf{L} = (L_{\alpha\beta})$ is symmetric. Thus the Onsager reciprocity relations are a consequence of the Gallavotti-Cohen symmetry.

Exercise 10.9 Consider a Brownian particle of mobility μ, immersed in a fluid at temperature T and subject to a force f parallel to the x axis. In this situation, the average displacement of the particle during a time interval of duration $\mathcal{T}$ is given by

$$\langle \Delta \boldsymbol{r} \rangle = \mu f \, \mathcal{T} \, \boldsymbol{i},$$

where $\boldsymbol{r} = (x, y, z)$ and $\boldsymbol{i}$ is the unit vector of the x axis. Show that the Jarzynski equality applied to this system implies the Einstein relation between the mobility and the diffusion coefficient D, that is,

$$D = k_\mathrm{B} T \, \mu.$$

Exercise 10.10 (Hamiltonian system) Consider a system satisfying the canonical equations of motion with state vector $x = ((q_1, p_1), \ldots, (q_n, p_n))$ and Hamiltonian $\mathcal{H}(x; \lambda) = \mathcal{H}(q_1, p_1, \ldots, q_n, p_n; \lambda)$. Assume that the system is in equilibrium at the temperature T for $t = t_0$, and is then manipulated according to a protocol $\boldsymbol{\lambda} = (\lambda(t))$ for $t_0 \le t \le t_\mathrm{f}$. During the manipulation, the system is isolated. Show that the Jarzynski equality holds for this system, where the work $\mathcal{W}(\boldsymbol{x}; \boldsymbol{\lambda})$ is defined in terms of the phase-space trajectory $\boldsymbol{x} = (x(t))$ and the protocol $\boldsymbol{\lambda}$ by

$$\mathcal{W}(\boldsymbol{x}; \boldsymbol{\lambda}) = \int_{t_0}^{t_\mathrm{f}} \mathrm{d}t \, \frac{\mathrm{d}\lambda(t)}{\mathrm{d}t} \, \frac{\partial \mathcal{H}(x(t); \lambda(t))}{\partial \lambda},$$

and ΔF is given by $F(\lambda_\mathrm{f}) - F(\lambda_0)$, with

$$F(\lambda) = -k_\mathrm{B} T \, \ln \int \mathrm{d}x \, \mathrm{e}^{-\mathcal{H}(x;\lambda)/k_\mathrm{B} T}.$$

Exercise 10.11 Show that the Crooks relation implies that the probability $P(\mathcal{W} \le \Delta F - \zeta)$ that the work $\mathcal{W}$ is smaller than the free energy difference ΔF by at least the amount $\zeta > 0$ satisfies the inequality

$$P(\mathcal{W} \le \Delta F - \zeta) \le \mathrm{e}^{-\zeta/k_\mathrm{B} T}.$$

Exercise 10.12 (Work as a sufficient statistics) Imagine that you are shown a movie of a mesoscopic system undergoing manipulation while in contact with a reservoir at temperature T. The movie can be either the record of the forward (F) process, in which the system is manipulated out of an initial equilibrium state, or the record of the backward (B) process, projected in reverse order, with equal a priori probabilities. Taking advantage

of the Crooks relation (10.63), show that to assess the probability that the movie shows the F process it is sufficient to evaluate the dissipated work $\mathcal{W} - \Delta F$, where the work $\mathcal{W}$ is computed from the trajectory exhibited in the movie. Evaluate the probability $p_{\mathrm{F}|\mathcal{W}}$ as a function of $\mathcal{W}$.

Exercise 10.13 (Jarzynski equality close to the thermodynamic limit) Consider a weakly interacting, dilute gas contained in a cylinder with a piston, initially at equilibrium at the temperature T. The system is then thermally isolated and undergoes an adiabatic quasi-static transformation in which its volume changes from V_0 to V_f. Assuming that transformation is performed infinitely slowly, verify the Jarzynski equality for this system. Where does stochasticity come from in this case?

Exercise 10.14 Show that the Widom insertion method for the evaluation of the chemical potential, discussed in an exercise in sec. 8.5, can be justified by the Jarzynski equality.

10.8 Adiabatic and Nonadiabatic Entropy Production

In the general case, where the system undergoes both manipulation and driving, the entropy production splits into two contributions, each nonnegative on average, where one vanishes when manipulation is absent, and the other when driving is absent. The first one is called the **adiabatic entropy production** and is defined for the trajectory $\boldsymbol{x}$ by

$$\Delta\mathcal{S}^{\mathrm{a}}[\boldsymbol{x}] = k_\mathrm{B} \sum_{k=1}^{n} \ln \frac{k_{x_k x_{k-1}}(t_k) p^{\mathrm{st}}_{x_{k-1}}(t_k)}{k_{x_{k-1} x_k}(t_k) p^{\mathrm{st}}_{x_k}(t_k)}. \tag{10.64}$$

Here $k_{xx'}(t) = k_{xx'}(\lambda(t))$ for a system manipulated according to the protocol $\boldsymbol{\lambda} = (\lambda(t))$. The second one is the **nonadiabatic entropy production**, defined by

$$\Delta\mathcal{S}^{\mathrm{na}}[\boldsymbol{x}] = k_\mathrm{B} \sum_{k=1}^{n} \ln \frac{p^{\mathrm{st}}_{x_k}(t_k)}{p^{\mathrm{st}}_{x_{k-1}}(t_k)} + k_\mathrm{B} \ln \frac{p_{x_0}(t_0)}{p_{x_\mathrm{f}}(t_\mathrm{f})}. \tag{10.65}$$

In both definitions, $p^{\mathrm{st}}(t)$ is the **instantaneous stationary distribution**, namely, the stationary probability distribution over x that would obtain if the transition rates $k_{xx'}(t)$ were frozen at the values they have at time t. We can check that the total entropy production $\Delta\mathcal{S}^{\mathrm{tot}}[\boldsymbol{x}]$ is given by the sum of $\Delta\mathcal{S}^{\mathrm{a}}[\boldsymbol{x}]$ and $\Delta\mathcal{S}^{\mathrm{na}}[\boldsymbol{x}]$.

Let us first consider the case of a system with jump rates constant in time, having settled of the corresponding nonequilibrium steady state described by p^{st}. It is then clear from eq. (10.64) that $\Delta\mathcal{S}^{\mathrm{a}}[\boldsymbol{x}]$ is the entropy produced to maintain the system in the nonequilibrium steady state p^{st}. On the other hand, $\Delta\mathcal{S}^{\mathrm{na}}[\boldsymbol{x}]$ vanishes in this situation. In the case in which no driving is present, namely when the jump rates satisfy the detailed balance condition (10.7), each term in eq. (10.64) vanishes, and we have $\Delta\mathcal{S}^{\mathrm{tot}}[\boldsymbol{x}] = \Delta\mathcal{S}^{\mathrm{na}}[\boldsymbol{x}]$.

Let us now suppose that the system is initially in a nonequilibrium steady state $p^{\mathrm{st}}(t_0)$ and is then manipulated to reach a final steady state $p^{\mathrm{st}}(t_\mathrm{f})$. Then the second term of

eq. (10.65) can be written $k_B(\ln p^{st}_{x_0}(t_0) - \ln p^{st}_{x_f}(t_f))$ and, by rearranging the terms, the associated nonadiabatic entropy production can be expressed by

$$\Delta \mathcal{S}^{na}[\boldsymbol{x}] = -k_B \sum_{k=0}^{n} \ln \frac{p^{st}_{x_k}(t_{k+1})}{p^{st}_{x_k}(t_k)} = \int_{t_0}^{t_f} dt\, \frac{d\lambda(t)}{dt}\, \frac{\partial \mathcal{S}^{st}_{x(t)}(\lambda(t))}{\partial \lambda}, \tag{10.66}$$

where

$$\mathcal{S}^{st}_x(\lambda) = -k_B \ln p^{st}_x(\lambda) \tag{10.67}$$

is the stochastic entropy associated with the instantaneous stationary distribution with the given value of λ. Since there is no contribution to the integral if $\lambda(t)$ remains constant from a given time $t_{end} < t_f$ onward, the nonadiabatic entropy production has the same expression even if we stop the integral at $t = t_{end}$, where the system is not yet in a steady state.

Both adiabatic and nonadiabatic entropy productions satisfy separate fluctuation relations. To obtain them, we introduce the **conjugate jump rates**

$$k^+_{xx'}(t) = k_{x'x}(t)\, \frac{p^{st}_x(t)}{p^{st}_{x'}(t)}. \tag{10.68}$$

These jump rates define a conjugate master equation, whose solution is $p^+(t) = (p^+_x(t))$. However, the stationary solutions of the original and of the conjugate master equation are the same, that is, $p^{st+}(t) = p^{st}(t)$. We have indeed

$$\sum_{x'} k^+_{xx'}(t) p^{st}_{x'}(t) = \sum_{x'} k_{x'x}(t) p^{st}_x(t) = \sum_{x'} k_{xx'}(t) p^{st}_{x'}(t) = \sum_{x'} k^+_{x'x} p^{st}_x(t). \tag{10.69}$$

The conjugate jump rates allow us to express the adiabatic and nonadiabatic entropy productions in terms of ratios of probabilities. Following the same steps as in sec. 10.6, we obtain

$$\Delta \mathcal{S}^{a}[\boldsymbol{x}; \boldsymbol{\lambda}] = k_B \ln \frac{\mathcal{P}[\boldsymbol{x}; \boldsymbol{\lambda}]}{\mathcal{P}^+[\boldsymbol{x}; \boldsymbol{\lambda}]}; \tag{10.70}$$

$$\Delta \mathcal{S}^{na}[\boldsymbol{x}; \boldsymbol{\lambda}] = k_B \ln \frac{\mathcal{P}[\boldsymbol{x}; \boldsymbol{\lambda}]}{\mathcal{P}^+[\widehat{\boldsymbol{x}}; \widehat{\boldsymbol{\lambda}}]}. \tag{10.71}$$

Here $\mathcal{P}^+[\boldsymbol{x}; \boldsymbol{\lambda}]$ is the probability density for trajectory $\boldsymbol{x}$ under the protocol $\boldsymbol{\lambda}$ under the conjugate master equation defined by the jump rates k^+. From these relations we obtain the integral fluctuation relations

$$\langle e^{-\Delta \mathcal{S}^a / k_B} \rangle = 1; \tag{10.72}$$

$$\langle e^{-\Delta \mathcal{S}^{na} / k_B} \rangle = 1. \tag{10.73}$$

By the inequality (10.54), we thus obtain that both adiabatic and nonadiabatic entropy productions are nonnegative on average. If the system is in a nonequilibrium steady state at the beginning of the manipulation, the nonadiabatic entropy production can be expressed by eq. (10.66), and we obtain

$$\left\langle \exp\left[-\frac{1}{k_B} \int_{t_0}^{t_f} dt\, \frac{d\lambda(t)}{dt}\, \frac{\partial \mathcal{S}^{st}_{x(t)}(\lambda(t))}{\partial \lambda} \right] \right\rangle = 1. \tag{10.74}$$

This is known as the **Hatano-Sasa relation**, and can be considered as the extension of the Jarzynski equality to nonequilibrium steady states.

Exercise 10.15 Derive the relations (10.70) and (10.71) explicitly.

10.9 Thermodynamics of Information

Information is physical.

—Rolf Landauer

The statistical nature of the second principle was illustrated by Maxwell with his famous "demon": he imagined a container holding a gas at a given temperature, divided in two sections, separated by a wall in which there is a passage that could be opened or closed by a "being" (called "demon" by Kelvin). The demon could let more energetic gas particles go, say, from right to left, and less energetic ones go from left to right. In this way, starting from equal temperatures in both compartments, the demon's actions would produce a temperature difference, that is, lower the entropy of the container. This thought experiment has been intensively discussed, and proposed in different disguises. One of the most popular ones is due to Szilard. He considers a cylindrical container of volume V, holding a gas made of a single molecule, in contact with a reservoir at the temperature T. The container can be divided into two sections of equal volume by a mobile partition acting as a piston, to which a shaft can be attached. Once the partition is inserted, the demon identifies the section that contains the particle, and attaches (in principle without expending work) the shaft on the opposite side of the partition. It can then allow the one-particle gas to expand reversibly until the gas occupies the whole volume V. In this way the demon is able to extract from the gas work given by

$$W = \int_{V/2}^{V} \mathrm{d}V'\, p(V') = \int_{V/2}^{V} \mathrm{d}V'\, \frac{k_\mathrm{B} T}{V'} = k_\mathrm{B} T \ln 2. \tag{10.75}$$

The partition is then removed and the system apparently returns to its initial state. We have thus obtained in principle a counterexample of the Kelvin statement of the second principle, according to which it is not possible to extract work with a cyclic process from a single heat reservoir.

Among the various attempts to reconcile this argument with the second principle, the most influential one is due to Landauer, who pointed out that the system does not exactly undergo a cyclic process. Indeed, at the beginning the demon must determine which side of the partition the particle is in, and it keeps this memory to the end. Landauer suggested therefore that work of size at least equal to $k_\mathrm{B} T \ln 2$ must be dissipated to erase the demon's memory, in order to maintain the validity of the second principle. This suggestion has been called "Landauer's principle" by Bennett, who enounces it as follows:

> Any *logically irreversible* manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding *entropy increase* in

non-information-bearing degrees of freedom of the information-processing apparatus or its environment.

It is then natural to consider that the system is made by an object, namely the cylinder, and a measuring device. In the simplest case, the system can be in one of two states, for example, $x \in \{\mathrm{L}, \mathrm{R}\}$ (for "left" or "right" of the partition), and the device in one of two analog states, denoted by y. Then the working of the Szilard engine can be described as follows:

1. The piston is initially placed in the middle of the cylinder, but no measurement is performed. Then x and y are independent variables that each assume the value L or R with probability 1/2. The system entropy is $\mathcal{S}^{\mathrm{sys}} = k_{\mathrm{B}} \ln 4 = 2k_{\mathrm{B}} \ln 2$.

2. After an error-free measurement, we have $(x, y) \in \{(\mathrm{L}, \mathrm{L}), (\mathrm{R}, \mathrm{R})\}$ with equal probability. Thus the system entropy is reduced to $k_{\mathrm{B}} \ln 2$. Therefore, either an amount of work equal to at least $k_{\mathrm{B}} T \ln 2$ has been made on the system and dissipated to the reservoir, or the entropy of a system attached to the device (for example, a memory) has increased by $k_{\mathrm{B}} T \ln 2$, or a combination of the two has occurred.

3. By performing the expansion, an amount of work equal to $k_{\mathrm{B}} T \ln 2$ is extracted from the reservoir.

4. The piston is placed again in the middle of the cylinder, and therefore the variables x and y are again independent.

The measurement in step 2 does not necessarily require dissipation. Indeed, if the device is initially in a reference state, for example, L, the measurement is a logically reversible operation: the state (L, L) maps to (L, L) and the state (R, L) maps to (R, R). On the other hand, at the end of the cycle, the device will be in state L or R with equal probability, and the dissipation of $k_{\mathrm{B}} T \ln 2$ will be necessary to bring it back to L (independently of the state it is in) according to Landauer's principle.

This discussion can be made more general and precise by exploiting fluctuation relations. Consider a system with states denoted by x that is manipulated according to a protocol. At a given time t_{m} an instantaneous (but not necessarily error-free) measurement is performed, and the outcome is denoted by y. The further protocol is chosen according to y. We are therefore implementing a **feedback protocol**, in which the manipulation depends on the state of the system. Thus the protocol $\boldsymbol{\lambda}$ is a function not only of the time t, but (for $t_{\mathrm{m}} < t \leq t_{\mathrm{f}}$) also of the measurement result y. The probability distribution of x and y immediately after the measurement is given by

$$p_{x,y}(t_{\mathrm{m}}) = p_{x|y}(t_{\mathrm{m}}) p_x(t_{\mathrm{m}}), \tag{10.76}$$

where $p_{x|y}(t)$ is the conditional probability of x at time t, given y. The mutual dependence of x and y immediately after the measurement is quantified by the **mutual information** I defined by (see Appendix F.2)

$$I(x:y) = \sum_{xy} p_{x,y}(t_{\mathrm{m}}) \ln \frac{p_{x,y}(t_{\mathrm{m}})}{p_x(t_{\mathrm{m}}) p_y(t_{\mathrm{m}})}. \tag{10.77}$$

This expression is the average of the **stochastic mutual information** $\mathcal{I}_{x;y}$ defined by

$$\mathcal{I}_{x;y} = \ln \frac{p_{x,y}(t_\mathrm{m})}{p_x(t_\mathrm{m}) p_y(t_\mathrm{m})} = \ln \frac{p_{x|y}(t_\mathrm{m})}{p_x(t_\mathrm{m})}. \tag{10.78}$$

The joint probability of a trajectory $\boldsymbol{x}$ of the object and a measurement outcome y is given by

$$\mathcal{P}_y[\boldsymbol{x};\boldsymbol{\lambda}] = p_{y|\boldsymbol{x}}(\boldsymbol{\lambda})\, \mathcal{P}[\boldsymbol{x};\boldsymbol{\lambda}]. \tag{10.79}$$

We can thus generalize the integral fluctuation relation in the presence of feedback control:

$$\begin{aligned} \left\langle \mathrm{e}^{-\Delta \mathcal{S}^\mathrm{tot}/k_\mathrm{B} - \mathcal{I}_{x;y}} \right\rangle &= \sum_y \int \mathcal{D}\boldsymbol{x}\, p_{y|\boldsymbol{x}} \mathcal{P}[\boldsymbol{x};\boldsymbol{\lambda}]\, \frac{\mathcal{P}[\widehat{\boldsymbol{x}};\widehat{\boldsymbol{\lambda}}]}{\mathcal{P}[\boldsymbol{x};\boldsymbol{\lambda}]}\, \frac{p_y(t_\mathrm{m})}{p_{y|\boldsymbol{x}}(t_\mathrm{m})} \\ &= \sum_y \int \mathcal{D}\boldsymbol{x}\, \mathcal{P}[\widehat{\boldsymbol{x}};\widehat{\boldsymbol{\lambda}}]\, p_y(t_\mathrm{m}) = 1. \end{aligned} \tag{10.80}$$

Here we made use of the fact that $p_{y|\boldsymbol{x}} = p_{y|\boldsymbol{x}}(t_\mathrm{m})$ since y depends only on the state of $\boldsymbol{x}$ at time t_m. This relation is known as the **Sagawa-Ueda relation**. It implies the inequality

$$\langle \Delta \mathcal{S}^\mathrm{tot} \rangle \geq -k_\mathrm{B} \left\langle \mathcal{I}_{x;y} \right\rangle = -k_\mathrm{B}\, I, \tag{10.81}$$

where I is the mutual information between $x(t_\mathrm{m})$ and the measurement outcome y.

As in the Jarzynski equality, the entropy production can be expressed in terms of the work $\mathcal{W}$ performed on the system if the protocol starts and ends in equilibrium. One has, however, to take into account that the free-energy difference ΔF may depend on the protocol, and therefore on the measurement outcome y. We obtain therefore the inequality

$$\langle \mathcal{W} \rangle - \langle \Delta F \rangle \geq -k_\mathrm{B} T\, I. \tag{10.82}$$

It is therefore possible to obtain useful work if the mutual information between $x(t_\mathrm{m})$ and y is significant, that is, if the measurement error is small enough.

Exercise 10.16 A 1-bit memory is represented by a system with three states: $x \in \{-1, 0, +1\}$, where $\epsilon_{\pm 1} = 0$ and $\epsilon_0 \gg k_\mathrm{B} T$, and only the jumps $0 \leftrightharpoons \pm 1$ are allowed, all with equal attempt frequency. The state 0 therefore acts as a barrier between the ± 1 states. The dynamics satisfies detailed balance. One can erase the contents of the memory by manipulating the energies according to the following protocol:

1. The energy barrier ϵ_0 is linearly lowered to 0.
2. Then the energies ϵ_0 and ϵ_{+1} are tilted by setting $\epsilon_0(t) = -\epsilon(t)/2$, $\epsilon_{+1}(t) = -\epsilon(t)$, where $\epsilon(t)$ grows linearly from 0 to a positive value ϵ_f. This biases the occupation toward the state $+1$.
3. The energy barrier ϵ_0 is linearly reset to its initial value.
4. Finally, the energy ϵ_{+1} is linearly reset to its initial value.

The outcome of the process is that, independently of the initial state, the final state of the system is equal to $+1$ with high probability. Evaluate W numerically for different values

of the duration $\mathcal{T}$ of the manipulation and check that the Landauer bound is approached as $\mathcal{T} \to \infty$. Also show numerically that the average work W in the reverse process, with the initial condition $p_x(t_0) = \delta_{x,+1}$, satisfies the inequality $W \geq -k_B T \ln 2$, and interpret this result.

10.10 Information Reservoirs

A thermal reservoir is characterized by the property that its entropy changes are proportional to the corresponding internal energy change:

$$T\,\Delta S = \Delta E, \tag{10.83}$$

where T is the absolute temperature of the reservoir. On the other hand, in a system such as a lifted weight, which acts as a mechanical energy reservoir, energy changes do not entail entropy changes. In view of this relation, we can describe such an energy reservoir as a thermodynamical system at zero temperature. We can also envisage an **information reservoir**, namely, system whose entropy can change without its internal energy being changed. Such a system can be made, for example, by a collection of microscopic objects that can each be in one of two states of equal energy, such as some crystal defects, or spins in an anisotropic paramagnet in zero field. Writing eq. (10.83) in the form

$$\Delta S = \frac{\Delta E}{T}, \tag{10.84}$$

we see that to such a system corresponds an infinite temperature.

Information reservoirs are often represented by an infinite tape containing cases, where each case contains a binary variable $y \in \{0, 1\}$, that is, one bit. If all cases are in the same state, for example, 1, the tape entropy vanishes. Otherwise, if all cases are independent, and the probability that $y = 1$ is equal to r, the tape entropy equals $S = k_B H_S(p^{(r)})$ per case, where $H_S(p^{(r)}) = -r \ln r - (1 - r) \ln(1 - r)$ is the Shannon entropy of the probability distribution $p^{(r)} = (p_0{=}1 - r, p_1{=}r)$. An idealized model of a system allowing the extraction of work from a heat reservoir by means of interaction with a tape can be described as follows. The system has two states, d (down) and u (up), with energies $\epsilon_d = 0$ and $\epsilon_u = \epsilon > 0$, and is in contact with a reservoir at the temperature T. The jump rates between d and u satisfy the detailed balance condition

$$\frac{k^{\uparrow}}{k^{\downarrow}} = \mathrm{e}^{-\epsilon/k_B T}, \tag{10.85}$$

where $k^{\uparrow} = k_{ud}$ and $k^{\downarrow} = k_{du}$. In the absence of other interactions, the equilibrium distribution of the system is

$$p_d^{\mathrm{eq}} = \frac{1}{1 + \mathrm{e}^{-\epsilon/k_B T}}; \qquad p_u^{\mathrm{eq}} = \frac{\mathrm{e}^{-\epsilon/k_B T}}{1 + \mathrm{e}^{-\epsilon/k_B T}}. \tag{10.86}$$

The system is also in contact with a tape. Interactions with the tape take place at random times, such that the probability of an interaction in a short interval of time of duration

dt is given by $\gamma\, dt$. In the interaction, the system can swap the two states. Namely, if the system is in u at the interaction time and the bit on the tape reads 1, the system state goes to d, the bit on the tape is set to 0, and an amount of energy ϵ is gleaned. Otherwise, if the system is in d and the bit on the tape is 0, the system goes to u and the bit is set to 1 and work equal to ϵ is performed on the system. In this way microscopic reversibility is maintained. The jump rates for this process are given by

$$\tilde{k}^{\uparrow} = \gamma\,(1-r); \qquad \tilde{k}^{\downarrow} = \gamma\, r. \tag{10.87}$$

Here r is the fraction of 1 bits in the incoming tape. After each interaction of this kind, the tape advances by one case and a fresh bit is exposed.

The evolution of the system is described by the master equation

$$\begin{aligned} \frac{dp_u}{dt} &= \left(k^{\uparrow} + \tilde{k}^{\uparrow}\right) p_d - \left(k^{\downarrow} + \tilde{k}^{\downarrow}\right) p_u; \\ \frac{dp_d}{dt} &= \left(k^{\downarrow} + \tilde{k}^{\downarrow}\right) p_u - \left(k^{\uparrow} + \tilde{k}^{\uparrow}\right) p_d. \end{aligned} \tag{10.88}$$

The system eventually reaches a steady state given by

$$p_u^{\mathrm{st}} = \frac{k^{\uparrow} + \tilde{k}^{\uparrow}}{k+\gamma}; \qquad p_d^{\mathrm{st}} = \frac{k^{\downarrow} + \tilde{k}^{\downarrow}}{k+\gamma}. \tag{10.89}$$

Here we have defined $k = k^{\uparrow} + k^{\downarrow}$ and we have used the relation $\gamma = \tilde{k}^{\uparrow} + \tilde{k}^{\downarrow}$. The average work extracted by the system per unit time is given by

$$\dot{W} = \epsilon\gamma \left[r p_u^{\mathrm{st}} - (1-r) p_d^{\mathrm{st}} \right]. \tag{10.90}$$

This work is positive, that is, the system is able to extract work, if

$$\frac{r}{1-r} > \frac{p_d^{\mathrm{st}}}{p_u^{\mathrm{st}}} = \frac{k^{\downarrow} + \gamma\, r}{k^{\uparrow} + \gamma\,(1-r)}. \tag{10.91}$$

The equality obtains for

$$\frac{r}{1-r} = \frac{k^{\downarrow}}{k^{\uparrow}} = \mathrm{e}^{\epsilon/k_{\mathrm{B}}T}. \tag{10.92}$$

Defining δ by the relation

$$r = \frac{\mathrm{e}^{\delta/k_{\mathrm{B}}T}}{1 + \mathrm{e}^{\delta/k_{\mathrm{B}}T}}, \tag{10.93}$$

eq. (10.91) assumes the simple form

$$\delta > \epsilon. \tag{10.94}$$

Since the fraction of 1 bits on the incoming tape is equal to r and that of the outgoing bits is equal to p_u^{st}, the average rate of the entropy change in the tape is given by

$$\dot{S}^{\mathrm{tape}} = k_{\mathrm{B}}\gamma \left[H_{\mathrm{S}}(p^{\mathrm{st}}) - H_{\mathrm{S}}(p^{(r)}) \right]. \tag{10.95}$$

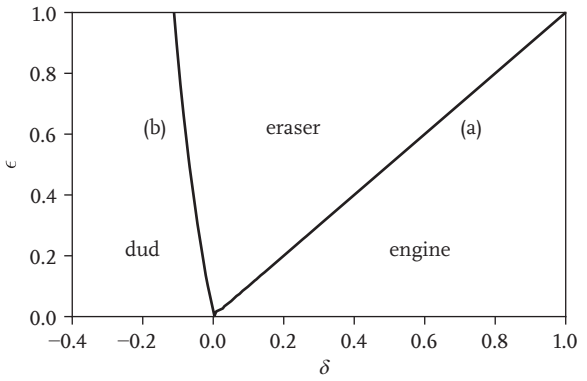


FIGURE 10.4. Phase diagram of the information-fueled engine in the (δ, ϵ) plane for $\gamma = 5\, k^{\downarrow}$. The line (a), $\delta = \epsilon$, separates the "engine" and the "eraser" regimes. The line (b), $H_S(p^{st}) = H_S(r)$, separates the "eraser" and the "dud" regimes. Energies are measured in units of $k_B T$.

If $\delta > \epsilon$, this expression is positive. On the other hand, if $\delta < \epsilon$, it may be positive or negative. If it is negative, the system acts as an eraser, that is, it expends work to reduce the entropy of the tape. There is also another regime, for $\delta < 0$ and large enough in absolute value, for which $H_S(p^{st}) > H_S(r)$. In this regime, the system neither performs useful work, nor reduces the entropy of the tape. It can be then described as a "dud". The resulting phase diagram is shown in fig. 10.4.

It is worth pointing out that the work spent in erasing the bits on the tape can be recovered in principle by feeding the outgoing tape to a similar engine (with the roles of $y = 1$ and 0 exchanged). From this point of view the Szilard engine and the Landauer principle are two sides of the same coin.

Exercise 10.17 Equation (10.88) can also be interpreted in the measurement-and-feedback scheme. In this interpretation, the information reservoir is a device that measures the state of the system at a rate γ, obtaining a result $y \in \{0, 1\}$, where 1 corresponds to u and 0 to d. The parameter r represents the error probability of the measurement, that is, $p_{u|1} = p_{d|0} = 1 - r$ and $p_{u|0} = p_{d|1} = r$. We assume $0 < r \leq 1/2$. If the outcome of the measurement is 1, the states are swapped. Evaluate the average work per unit time performed by the system and the total entropy production in this interpretation and connect these quantities to the mutual entropy of the measurement via the Sagawa-Ueda inequality (10.81).

Exercise 10.18 (System interacting with two tapes) Consider a two-state system with two edges between the states, labeled I and II. Both edges are associated with information-mediated jumps, driven by interaction with tapes. The fractions of 1s in the two incoming tapes are $r_I \leq \frac{1}{2}$ and $r_{II} \leq \frac{1}{2}$. The jump rates from 0 to 1 are $\gamma_I r_I$ for edge I and $\gamma_{II} r_{II}$ for edge II. The reverse jump rates are $\gamma_I(1 - r_I)$ and $\gamma_{II}(1 - r_{II})$.

Evaluate the entropy production rate. Assuming $r_I < r_{II} \leq \frac{1}{2}$, information is written on tape I and erased from tape II. Evaluate the efficiency η of the process, defined by

$$\eta = -\frac{\dot{S}^{\rm erase}}{\dot{S}^{\rm write}},$$

in the regimes $\gamma_{\rm II} \gg \gamma_{\rm I}$ and $\gamma_{\rm II} \ll \gamma_{\rm I}$, respectively.

10.11 Copying Information

In biological systems it is extremely important to copy information from one form into another. Examples are the transcription of DNA sequences into RNA, or the translation of mRNA into proteins. Information copying does not necessarily involve dissipation, as we have seen before when discussing measurement. However, dissipationless copying can only be achieved at vanishing speed. Moreover, copying errors are inevitable, due to thermal fluctuations. Correcting these errors can significantly increase dissipation. Understanding the trade-offs between error rate, speed, and dissipation is important for understanding the functioning of biological machines.

Let us dwell on the example of a mesoscopic machine that sequentially replicates a long chain polymer. To keep things simple, we assume that the polymer is made up of two kinds of monomers. At each instant, the monomer at the tip of the copied polymer can be removed, or a new monomer can be added. We denote by $k_{\rm r}^+$, $k_{\rm r}^-$ the rates for incorporation and removal of the right monomer, and by $k_{\rm w}^+$, $k_{\rm w}^-$ the corresponding rates for the wrong monomer. The generalized detailed balance condition (10.17) then reads

$$\frac{k_{\rm r}^+}{k_{\rm r}^-} = {\rm e}^{(-\Delta\epsilon_{\rm r}+q)/k_{\rm B}T}; \qquad \frac{k_{\rm w}^+}{k_{\rm w}^-} = {\rm e}^{(-\Delta\epsilon_{\rm w}+q)/k_{\rm B}T}. \tag{10.96}$$

Here $\Delta\epsilon_{\rm r}$ and $\Delta\epsilon_{\rm w}$ are, respectively, the energies of the incorporation of a right and a wrong monomer, and q is the external driving that promotes monomer incorporation. We assume for simplicity that the driving is equal for right or wrong monomers. The average work W per incorporated monomer is given by

$$W = q. \tag{10.97}$$

Increasing q drives the system further and further away from equilibrium. We can thus write the jump rates as follows:

$$k_{\rm r}^+ = \omega_{\rm r}\, {\rm e}^{q/k_{\rm B}T}; \quad k_{\rm r}^- = \omega_{\rm r}\, {\rm e}^{\Delta\epsilon_{\rm r}/k_{\rm B}T}; \tag{10.98}$$

$$k_{\rm w}^+ = \omega_{\rm w}\, {\rm e}^{q/k_{\rm B}T}; \quad k_{\rm r}^- = \omega_{\rm w}\, {\rm e}^{\Delta\epsilon_{\rm w}/k_{\rm B}T}. \tag{10.99}$$

We can see that the discrimination between r and w depends both on the ratio of the attempt frequencies $\omega_{\rm r,w}$ and on the difference of the binding energies $\Delta\epsilon_{\rm r,w}$.

We describe template copying by a three-state system, in which the state $\emptyset$ denotes the unbound state before the enzyme binds to a residue, and r and w, in which a right or a wrong residue, respectively, is attached to the copy. We denote by η the average fraction of wrong incorporations in the growing polymer. Then the average elongation speed of the

polymer is given by

$$v = k_{\rm r}^{+} - (1-\eta)k_{\rm r}^{-} + k_{\rm w}^{+} - \eta k_{\rm w}^{-}. \tag{10.100}$$

This implies

$$\frac{\eta}{1-\eta} = \frac{k_{\rm w}^{+} - \eta k_{\rm w}^{-}}{k_{\rm r}^{+} - (1-\eta)k_{\rm r}^{-}}. \tag{10.101}$$

In the steady state, one has

$$\frac{p_{\rm w}}{p_{\emptyset}} = \eta; \qquad \frac{p_{\rm r}}{p_{\emptyset}} = 1-\eta. \tag{10.102}$$

In the equilibrium limit, the elongation speed v goes to zero. We can then solve eq. (10.100) for η and substitute it in eq. (10.101), obtaining

$$\eta^{\rm eq} = {\rm e}^{-\Delta\epsilon_{\rm w}/k_{\rm B}T}/({\rm e}^{-\Delta\epsilon_{\rm w}/k_{\rm B}T} + {\rm e}^{-\Delta\epsilon_{\rm r}/k_{\rm B}T}), \tag{10.103}$$

as expected from detailed balance. The corresponding driving q is smaller than the average work needed to incorporate one monomer: $q^{\rm stall} = \epsilon_{\rm r} + k_{\rm B}T\ln(1-\eta^{\rm eq})$. The second (negative) term is due to the compositional disorder of the growing polymer. On the other hand, in the fully irreversible limit, $q \gg k_{\rm B}T$, the error rate approaches $k_{\rm w}^{+}/(k_{\rm w}^{+}+k_{\rm r}^{+}) = \omega_{\rm w}/(\omega_{\rm w}+\omega_{\rm r})$. In this limit, the discrimination is due only to the ratio of the attempt frequencies for incorporation.

The total entropy production rate is then given by the Schnakenberg formula (10.47):

$$\begin{aligned}
T\,\dot{S}^{\rm tot} &= k_{\rm B}T\left[J_{\rm w\emptyset}\ln\frac{k_{\rm w}^{+}p_{\emptyset}}{k_{\rm w}^{-}p_{\rm w}} + J_{\rm r\emptyset}\ln\frac{k_{\rm r}^{+}p_{\emptyset}}{k_{\rm r}^{-}p_{\rm r}}\right] \\
&= \eta\, v\, k_{\rm B}T\ln\frac{k_{\rm w}^{+}p_{\emptyset}}{k_{\rm w}^{-}p_{\rm w}} + (1-\eta)\, v\, k_{\rm B}T\ln\frac{k_{\rm r}^{+}p_{\emptyset}}{k_{\rm r}^{-}p_{\rm r}} \\
&= \eta\, v\, k_{\rm B}T\ln\frac{k_{\rm w}^{+}}{k_{\rm w}^{-}\eta} + (1-\eta)\, v\, k_{\rm B}T\ln\frac{k_{\rm r}^{+}}{k_{\rm r}^{-}(1-\eta)} \\
&= v\left[\eta\left(W - \Delta F + k_{\rm B}T\ln\frac{\eta^{\rm eq}}{\eta}\right)\right. \\
&\qquad \left. + (1-\eta)\left(W - \Delta F + k_{\rm B}T\ln\frac{1-\eta^{\rm eq}}{1-\eta}\right)\right] \\
&= v\left[W - \Delta F - k_{\rm B}T D_{\rm KL}(\eta\|\eta^{\rm eq})\right],
\end{aligned} \tag{10.104}$$

where $W = q$ is the average work per incorporated monomer and

$$\Delta F = k_{\rm B}T\,\ln\left({\rm e}^{-\Delta\epsilon_{\rm r}/k_{\rm B}T} + {\rm e}^{-\Delta\epsilon_{\rm r}/k_{\rm B}T}\right) \tag{10.105}$$

is the equilibrium free energy change upon incorporation of a monomer.

To interpret this result, consider the case in which there is no binding energy difference between right and wrong monomers, that is, $\eta^{\rm eq} = 1/2$. Also assume that it is possible to change reversibly one right monomer into a wrong one and back. Then the system made

by the template and the copy is conceptually equivalent to the tape system discussed in sec. 10.10. In particular, a precise match of the template and the copy can be fed to such a machine operating in the "engine" regime. This machine would then be able to extract work by increasing the error in the copy. Therefore the last term in eq. (10.104) represents free energy reversibly incorporated in the copy as information.

10.12 Uncertainty Relations

Uncertainty relations are universal inequalities relating the average value of an arbitrary current, its variance, and the average entropy production rate. In a nutshell, uncertainty relations state that achieving very accurate currents, where the root mean square of the deviation from the mean is much smaller than the mean, requires in general a minimal cost in terms of entropy production. In fact, a number of several interconnected results have been obtained, all broadly termed "uncertainty relations." They differ in the underlying hypotheses and provide tighter or looser bounds in different situations.

Here we confine our discussion to a thermodynamic uncertainty relation that holds in a nonequilibrium steady state and has the advantage of being relatively easy to prove.

We consider a system obeying a master equation (10.1) in the steady state $p^{\mathrm{st}} = (p_x^{\mathrm{st}})$. We focus on the joint probability distribution $p(\mathcal{S}^{\mathrm{tot}}, \mathcal{J})$ of the total entropy production $\mathcal{S}^{\mathrm{tot}}$ and a given empirical current $\mathcal{J}$, odd under time reversal, both integrated over the time interval of duration $\mathcal{T}$. The joint probability distribution satisfies a detailed fluctuation relation:

$$p(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}) = \mathrm{e}^{\mathcal{S}^{\mathrm{tot}}/k_{\mathrm{B}}}\, p(-\mathcal{S}^{\mathrm{tot}}, -\mathcal{J}). \tag{10.106}$$

We can prove eq. (10.106) as follows:

$$\begin{aligned} p(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}) &= \int \mathcal{D}\boldsymbol{x}\, \delta(\mathcal{S}^{\mathrm{tot}} - \mathcal{S}^{\mathrm{tot}}(\boldsymbol{x}))\, \delta(\mathcal{J} - \mathcal{J}(\boldsymbol{x}))\, \mathcal{P}(\boldsymbol{x}) \\ &= \int \mathcal{D}\boldsymbol{x}\, \delta(\mathcal{S}^{\mathrm{tot}} - \mathcal{S}^{\mathrm{tot}}(\boldsymbol{x}))\, \delta(\mathcal{J} - \mathcal{J}(\boldsymbol{x}))\, \mathrm{e}^{\mathcal{S}^{\mathrm{tot}}(x)}\, \mathcal{P}(\hat{\boldsymbol{x}}) \\ &= \mathrm{e}^{\mathcal{S}^{\mathrm{tot}}/k_{\mathrm{B}}} \int \mathcal{D}\hat{\boldsymbol{x}}\, \delta(\mathcal{S}^{\mathrm{tot}} + \mathcal{S}^{\mathrm{tot}}(\hat{\boldsymbol{x}}))\, \delta(\mathcal{J} + \mathcal{J}(\hat{\boldsymbol{x}}))\, \mathcal{P}(\hat{\boldsymbol{x}}) \\ &= \mathrm{e}^{\mathcal{S}^{\mathrm{tot}}/k_{\mathrm{B}}}\, p(-\mathcal{S}^{\mathrm{tot}}, -\mathcal{J}). \end{aligned} \tag{10.107}$$

We introduce a probability density normalized over the *positive* values of the total entropy production:

$$p^{+}(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}) = \left(1 + \mathrm{e}^{-\mathcal{S}^{\mathrm{tot}}/k_{\mathrm{B}}}\right)\, p(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}). \tag{10.108}$$

One has, by eq. (10.106),

$$\int_0^{\infty} \mathrm{d}\mathcal{S}^{\mathrm{tot}} \int_{-\infty}^{+\infty} \mathrm{d}\mathcal{J}\, p^{+}(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}) = \int_{-\infty}^{\infty} \mathrm{d}\mathcal{S}^{\mathrm{tot}} \int_{-\infty}^{+\infty} \mathrm{d}\mathcal{J}\, p(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}) = 1. \tag{10.109}$$

We then express the average of our integrated current as

$$\begin{aligned}\langle \mathcal{J} \rangle &= \int_{-\infty}^{+\infty} \mathrm{d}\mathcal{S}^{\mathrm{tot}} \int_{-\infty}^{+\infty} \mathrm{d}\mathcal{J}\, \mathcal{J}\, p(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}) \\ &= \int_{0}^{\infty} \mathrm{d}\mathcal{S}^{\mathrm{tot}} \int_{-\infty}^{+\infty} \mathrm{d}\mathcal{J}\, \mathcal{J}\, p(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}) \left(1 - \mathrm{e}^{-\mathcal{S}^{\mathrm{tot}}/k_{\mathrm{B}}}\right) \\ &= \int_{0}^{\infty} \mathrm{d}\mathcal{S}^{\mathrm{tot}} \int_{-\infty}^{+\infty} \mathrm{d}\mathcal{J}\, \mathcal{J}\, p^{+}(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}) \frac{1 - \mathrm{e}^{-\mathcal{S}^{\mathrm{tot}}/k_{\mathrm{B}}}}{1 + \mathrm{e}^{-\mathcal{S}^{\mathrm{tot}}/k_{\mathrm{B}}}} \\ &= \left\langle \mathcal{J} \tanh\left(\frac{\mathcal{S}^{\mathrm{tot}}}{2k_{\mathrm{B}}}\right) \right\rangle^{+},\end{aligned} \tag{10.110}$$

where we have defined

$$\langle \cdots \rangle^{+} = \int_{0}^{\infty} \mathrm{d}\mathcal{S}^{\mathrm{tot}} \int_{-\infty}^{+\infty} \mathrm{d}\mathcal{J}\, \cdots\, p^{+}(\mathcal{S}^{\mathrm{tot}}, \mathcal{J}). \tag{10.111}$$

The average entropy production can be similarly related to the average over positive values only:

$$\langle \mathcal{S}^{\mathrm{tot}} \rangle = \left\langle \mathcal{S}^{\mathrm{tot}} \tanh\left(\frac{\mathcal{S}^{\mathrm{tot}}}{2k_{\mathrm{B}}}\right) \right\rangle^{+}. \tag{10.112}$$

For any two real vectors u and v for which a scalar product $u \cdot v$ is defined, the Cauchy-Schwarz inequality holds:

$$(u \cdot v)^2 \leq u^2 v^2. \tag{10.113}$$

To prove it, observe that

$$(u + \lambda v)^2 = u^2 + 2\lambda (u \cdot v) + \lambda^2 v^2 \geq 0, \qquad \forall \lambda \in \mathbb{R}. \tag{10.114}$$

Therefore the discriminant $4\left((u \cdot v)^2 - u^2 v^2\right)$ of this quadratic equation in λ should not be positive, and the inequality follows. Applying this inequality to eq. (10.110) we obtain

$$\langle \mathcal{J} \rangle^2 = \left(\left\langle \mathcal{J} \tanh\left(\frac{\mathcal{S}^{\mathrm{tot}}}{2k_{\mathrm{B}}}\right) \right\rangle^{+} \right)^2 \leq \langle \mathcal{J}^2 \rangle^{+} \left\langle \tanh^2\left(\frac{\mathcal{S}^{\mathrm{tot}}}{2k_{\mathrm{B}}}\right) \right\rangle^{+}. \tag{10.115}$$

For any nonnegative variable x, one has

$$\left\langle \tanh^2 x \right\rangle \leq \tanh \langle x \rangle . \tag{10.116}$$

Indeed, if $x \geq 0$, one has $\tanh^2 x \leq \tanh x$, and moreover $\left\langle \tanh x \right\rangle \leq \tanh \langle x \rangle$ by the Jensen inequality, since $\tanh x$ is concave for non-negative x. We obtain therefore the following

thermodynamic uncertainty relation:

$$\frac{\sigma_{\mathcal{J}}^2}{\langle \mathcal{J} \rangle^2} \geq \frac{2}{e^{S^{\rm tot}/k_{\rm B}} - 1}, \tag{10.117}$$

where $S^{\rm tot} = \langle \mathcal{S}^{\rm tot} \rangle$ and $\sigma_{\mathcal{J}}^2 = \langle \mathcal{J}^2 \rangle - \langle \mathcal{J} \rangle^2$ and we use that $\langle \mathcal{J}^2 \rangle = \langle \mathcal{J}^2 \rangle^+$.

The average of $\mathcal{J}$ is proportional to the duration $\mathcal{T}$ of the time interval: $\langle \mathcal{J} \rangle = \mathcal{T} J^{\rm st}$, where $J^{\rm st}$ is the steady-state average of the associated intensive current. The same holds for the total entropy production: $\langle \mathcal{S}^{\rm tot} \rangle = \mathcal{T} \dot{S}^{\rm tot}$. Also, the variance of $\mathcal{J}$ is extensive, $\sigma_{\mathcal{J}}^2 = \tilde{\sigma}_J^2 \mathcal{T}$, which defines $\tilde{\sigma}_J^2$. Substituting in (10.117) and taking the small $\mathcal{T}$ limit we obtain

$$\frac{\tilde{\sigma}_J^2}{(J^{\rm st})^2} \geq \frac{2\, k_{\rm B}}{\dot{S}^{\rm tot}}, \qquad \mathcal{T} \ll k_{\rm B}/\dot{S}^{\rm tot}, \tag{10.118}$$

where we expanded $e^{S^{\rm tot}/k_{\rm B}} \approx 1 + \mathcal{T} \dot{S}^{\rm tot}/k_{\rm B}$ for small $\mathcal{T}$. Equation (10.118) implies that the fluctuations of the current $\mathcal{J}$, when its average is fixed, can be reduced only at the expense of an increased entropy production rate. It can actually be shown that eq. (10.118) holds for any $\mathcal{T}$.

We now use eq. (10.117) to bound the efficiency of a molecular motor. The molecular motor moves with an average speed $v = \langle \dot{r} \rangle$, where r is its position along a microtubule. The motor carries a load that exerts on it a force f. The average entropy production rate is given by

$$T\, \dot{S}^{\rm tot} = \dot{W}^{\rm chem} - f\, v, \tag{10.119}$$

where $W^{\rm chem}$ is the average rate of consumption of the chemical free energy. We express $W^{\rm chem}$ in terms of the average ATP consumption rate $J_n = \langle \dot{n} \rangle$ and the ATP chemical potential imbalance $\Delta\mu$ by

$$\dot{W}^{\rm chem} = \Delta\mu\ \langle \dot{n} \rangle\,. \tag{10.120}$$

The macroscopic efficiency of the motor is

$$\eta_S = \frac{f v}{\dot{W}^{\rm chem}} = \frac{f v}{f v + T\, \dot{S}^{\rm tot}}. \tag{10.121}$$

By substituting (10.118) we obtain

$$\eta_S \leq \frac{1}{1 + 2 v\, k_{\rm B} T/(\tilde{\sigma}_J^2 f)}. \tag{10.122}$$

The bound provided by eq. (10.122) is particularly useful since all quantities appearing in the right-hand side are in principle accessible to experiments.

Recommended Reading

Stochastic thermodynamics has by now a vast literature. Stochastic energetics is discussed in the book by K. Sekimoto, *Stochastic Energetics*, Berlin: Springer, 2010. Reviews on

stochastic thermodynamics are given by U. Seifert, "Stochastic thermodynamics, fluctuation theorems, and molecular machines," *Reports on Progress in Physics* **75** 126001 (2012) and "Stochastic thermodynamics: From principles to the cost of precision," *Physica A: Statistical Mechanics and Its Applications* **504** 176–191 (2018). See also C. Jarzynski, "Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale," *Annual Reviews of Condensed Matter Physics* **2** 329–351 (2011). A comprehensive introduction is the book by L. Peliti and S. Pigolotti, *Stochastic Thermodynamics: An Introduction*, Princeton: Princeton University Press, 2021.

11 Complex Systems

Every rule has exceptions, including this one.

—Anonymous

11.1 Introduction

In this chapter, we discuss systems in statistical mechanics that are characterized by disorder—not the usual thermal disorder, represented by the Boltzmann-Gibbs distribution, but a "frozen" disorder, in which the Hamiltonian itself is random. One usually speaks of **quenched** disorder, when the disorder configuration does not change over the observation time, constrasting it with the **annealed** disorder, where the disorder variables change over the observation time, and can thus be dealt with by the usual methods of statistical mechanics.

The study of these systems is obviously of great importance for the applications—one needs only to remember that all metal alloys are composed of mixtures of different metals, smelted together, and then cooled. In these systems, at room temperature, the arrangement of atoms of different chemical nature does not change even over very long time periods, and it is therefore set by the initial conditions. The semiconductors used in electronics also generally exhibit impurities (*doping*) distributed in a random fashion, and these remain blocked inside the solid matrix at room temperature. It is therefore interesting to determine how the presence of frozen disorder changes the behavior of a system.

Indeed, the study of statistical mechanics of systems with a random Hamiltonian has opened a new field of study. Among other applications, it is now possible to consider problems of a strictly geometric nature by means of a clever generalization of the statistical approach. An important case is that of **percolation**, which describes the behavior, say, of a lattice in which only a fraction of sites are occupied at random. When the fraction of occupied nodes is small, one has small disconnected clusters of occupied sites. At higher occupation fractions, these clusters merge into significantly larger clusters, which span the sample from one boundary to the opposite one. We thus have two regimes: at a low occupation fraction, an occupied site picked at random belongs almost certainly to a finite cluster. If the occupation fraction is large enough, there is a finite probability that such a

point belongs to a cluster that spans the sample. The occupation fraction that separates these two regimes is the **percolation threshold**.

A class of disordered systems whose study requires more sophisticated methods comprises those exhibiting **frustration**, whose paradigmatic example is given by **spin glasses**. In frustrated systems, there is no configuration such that each interaction term in the Hamiltonian has its smallest value. As a consequence, the microscopic states with the lowest values of energy are numerous but quite different from one another, and are often separated by high energy barriers. Thus, below a certain temperature, the system finds itself trapped in a region of its phase space determined by its initial condition. As a consequence, once the system is found to be in one of these states, it will never be able to jump into another one in a finite time. In this sense, the observable state is not the Boltzmann-Gibbs one, but one of these **pure states**. This looks analogous to spontaneous symmetry breaking, but differs from it since, in general, there is no transformation that maps one of these states to another, and because the free energy of these states is not the same for all of them, but is a random quantity with a nontrivial distribution.

Some innovative techniques have been introduced to allow the representation of systems with quenched disorder. The best known of these techniques is the **replica method**. It has produced surprising results, at least for those few models that can be treated exactly. The extent to which these results are valid for the real world is the object of intense debate. In this situation, numerical simulations perform an essential role because they allow one to obtain detailed information about the behavior of realistic, but analytically intractable, systems—information that is not experimentally accessible.

We start our discussion of disordered systems by dealing with the problem of linear polymers in a solvent, although, strictly speaking, quenched disorder is not present in this system. It is, however, a useful starting point, because it allows us to see how a strictly geometrical problem (that is what are the geometrical properties of random walks without intersections?) can be treated as the limit of an ordinary problem in statistical mechanics.

11.2 Linear Polymers in Solution

Polymers are molecules obtained by connecting a large number of simple units (the **monomers**) using covalent bonds. In **linear** polymers, the monomers are arranged in a sequence. Biologically important examples are proteins, in which the monomers are amino acids, connected to one another by means of the peptide bond $-NH-CO-$, and nucleic acids (DNA and RNA), whose units are nucleotides (molecules formed by a base—adenine, guanine, cytosine, and thymine or (uracil)—bonded with a sugar—deoxyribose or ribose—which in turn is bonded to a phosphate) that are connected by means of a phosphodiester bond $-O-P-O-$. Plastic materials are also usually made of linear polymers—polyethylene, for instance, has the structure $\cdots -CH_2-CH_2-\cdots$, which is a result of the polymerization of ethylene $CH_2=CH_2$.

We must therefore distinguish between **homogeneous** polymers such as polyethylene, in which the monomers all belong to one chemical species, or to several ones that succeed one another in a regular fashion, and **heterogeneous** polymers such as proteins and

nucleic acids, in which the monomers do not follow one another with regularity along the molecule. In the remainder of this section, we deal only with homogeneous linear polymers.

In these structures, the bond between successive monomers has a certain flexibility and allows them to rotate fairly freely around the axis that connects them. In this manner, the polymer, as soon as it is composed of a large enough number of monomers, can be treated like a flexible chain.

Linear polymers are usually produced in a **good solvent**—a solvent whose affinity for the monomers is greater than the affinity the monomers have for one another. In this situation, the solvent interposes itself between monomers and tends to keep the monomers apart. In a bad solvent, on the other hand, the solvent is chased off by the polymer, which therefore folds in on itself and tends to form a compact ball and to precipitate from the solution. The two regimes are separated by the **coil-globule transition**, which is also known as the **Θ-point**.

11.2.1 The Gaussian Model

Let us consider a linear isolated polymer in a good solvent, and ask ourselves what configurations it might assume. We denote the distance between the center of one monomer and the center of the next by a_0, and the **degree of polymerization**—that is, the number of monomers that make up the polymer—by N. Let us assume that the bond's structure allows three successive monomers to assume any configuration, as long as the distance between their respective centers remains equal to a_0.

If we temporarily disregard the **excluded volume interaction**—in other words, the significant repulsion exerted by monomers close to one another in ambient space—we obtain a model of the linear polymer called the **Rayleigh chain**. This model can be formally defined as follows:

1. The configuration of a polymer composed of $N+1$ monomers (and by N covalent bonds between one polymer and the next) is represented by $N+1$ variables $\boldsymbol{r}_i$, $i=0,\dots,N$, where $\boldsymbol{r}_i \in \mathbb{R}^3$ is the position of the center of the $(i+1)$th monomer.

2. The covalent bond that connects successive monomers is represented by the constraint $|\boldsymbol{r}_{i+1}-\boldsymbol{r}_i| = a_0$, $i=0,\dots,N$.

3. All the configurations that satisfy the constraint have the same energy (which can be set to zero).

The polymer will assume the shape of a random coil of approximate spherical shape, occupying a certain region of space. We want to calculate the typical linear dimensions of this region. We can first ask ourselves what is the characteristic distance between the first monomer (which we can assume placed at the origin: $\boldsymbol{r}_0 = 0$) and the last one, placed at $\boldsymbol{r}_N$. We can easily answer this question by observing that, in our hypotheses, the vectors

$$\boldsymbol{y}_i = \boldsymbol{r}_i - \boldsymbol{r}_{i-1}, \qquad i=1,\dots,N, \tag{11.1}$$

which represent the distance between successive monomers, are independent, identically distributed random vectors. Their distribution is uniform over a sphere of radius a_0,

$$P(\boldsymbol{\gamma}) \propto \delta\left(\gamma^2 - a_0^2\right), \tag{11.2}$$

where the proportionality constant is determined by the normalization condition

$$\int \mathrm{d}^3\boldsymbol{\gamma}\, P(\boldsymbol{\gamma}) = 1. \tag{11.3}$$

It is then easy to see that $\langle\boldsymbol{\gamma}\rangle = 0$ and $\langle\gamma^2\rangle = a_0^2$. Then, since $\boldsymbol{r}_0 = 0$, one has

$$\boldsymbol{r}_N = \sum_{i=1}^{N} \boldsymbol{\gamma}_i, \tag{11.4}$$

from which it follows $\langle \boldsymbol{r}_N \rangle = 0$, and

$$\left\langle r_N^2 \right\rangle = N\, a_0^2. \tag{11.5}$$

The coil's radius therefore increases as $N^{1/2}$. In order to have a more easily accessible measure of the coil size, one usually considers the **radius of gyration** R_G, which is defined by

$$R_\mathrm{G}^2 = \frac{1}{N+1} \sum_i |\boldsymbol{r}_i - \boldsymbol{R}_\mathrm{CM}|^2, \tag{11.6}$$

where $\boldsymbol{R}_\mathrm{CM} = \sum_i \boldsymbol{r}_i/(N+1)$ is the center of mass of the polymer. Then the moment of inertia of the coil is given by its mass times R_G^2.

Exercise 11.1 Evaluate $\left\langle R_\mathrm{G}^2 \right\rangle$ and its relation with $\left\langle r_N^2 \right\rangle$ in the Gaussian model of a linear polymer.

Because of eq. (11.4), where the $\boldsymbol{\gamma}_i$ are independent, identically distributed random vectors, with a vanishing average and variance equal to a_0^2, we can conclude from the central limit theorem (sec. A.6) that $\boldsymbol{r}_N$ is, for a sufficiently large N, a random Gaussian vector, with a vanishing average and variance equal to Na_0^2,

$$P_N(\boldsymbol{r}_N) \propto \exp\left(-\frac{r_N^2}{2N\, a_0^2}\right), \tag{11.7}$$

and this for any reasonable distribution of the $\boldsymbol{\gamma}$s. Since this relation necessarily follows from our hypothesis that it is possible to neglect of the excluded-volume interaction, the linear polymer model in which this hypothesis is made is called the **Gaussian model**.

We can now formulate the Gaussian model in a form that emphasizes its analogy with phase transitions. First, we generalize the problem to d dimensions. Then, we introduce $\Gamma_N(\boldsymbol{r})$, which counts the number of configurations in which $\boldsymbol{r}_0 = 0$ and $\boldsymbol{r}_N = \boldsymbol{r}$:

$$\Gamma_N(\boldsymbol{r}) = \int \prod_{i=1}^{N} \left\{\mathrm{d}\boldsymbol{r}_i\, \delta(|\boldsymbol{r}_i - \boldsymbol{r}_{i-1}|^2 - a_0^2)\right\}\, \delta\left(\boldsymbol{r}_N - \boldsymbol{r}\right). \tag{11.8}$$

We then obviously have

$$\Gamma_{N+1}(\boldsymbol{r}) = \mathcal{S} \int \mathrm{d}\boldsymbol{r}_N \, P(\boldsymbol{r} - \boldsymbol{r}_N) \, \Gamma_N(\boldsymbol{r}_N), \tag{11.9}$$

where $P(\boldsymbol{\gamma})$ is the probability distribution of the distance between successive monomers, and $\mathcal{S}$ is a constant that depends on the details of the distribution of the $\boldsymbol{\gamma}$.

This equation can be solved in Fourier space. Define

$$\Gamma_{N+1}(\boldsymbol{k}) = \int \mathrm{d}\boldsymbol{r} \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \, \Gamma_{N+1}(\boldsymbol{r}). \tag{11.10}$$

Since the right-hand side of eq. (11.9) is a convolution, we have (see sec. D.2)

$$\Gamma_{N+1}(\boldsymbol{k}) = \frac{\mathcal{S}}{(2\pi)^d} P(\boldsymbol{k}) \Gamma_N(\boldsymbol{k}). \tag{11.11}$$

Setting $\Gamma_0(\boldsymbol{r}) = \delta(\boldsymbol{r})$, we obtain

$$\Gamma_N(\boldsymbol{k}) = \left[\frac{\mathcal{S}}{(2\pi)^d} \right]^N P(\boldsymbol{k})^N. \tag{11.12}$$

Now $P(\boldsymbol{k})$ is the characteristic function of $P(\boldsymbol{r})$ (see sec. A.5). Therefore, by eq. (A.31), if $P(\boldsymbol{r})$ is such that $\langle \boldsymbol{r} \rangle = 0$ and $\langle r^2 \rangle = a_0^2$, then, for small values of $|\boldsymbol{k}|$, one has

$$P(\boldsymbol{k}) \simeq \mathrm{e}^{-a_0^2 k^2/2}. \tag{11.13}$$

Therefore

$$\Gamma_N(\boldsymbol{k}) \simeq \left[\frac{\mathcal{S}}{(2\pi)^d} \right]^N \mathrm{e}^{-N a_0^2 k^2/2}. \tag{11.14}$$

Taking the antitransform of $\Gamma_N(\boldsymbol{k})$, we obtain $\Gamma_N(\boldsymbol{r})$:

$$\Gamma_N(\boldsymbol{r}) = \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^d} \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \, \Gamma_N(\boldsymbol{k}) \sim \mathrm{e}^{-r^2/(2N a_0^2)}, \tag{11.15}$$

which obviously holds only for $r^2 \gg a_0^2$, so that the integral in (11.15) is dominated by small values of $|\boldsymbol{k}|$.

Let us now go over to the grand canonical ensemble, by introducing the fugacity z:

$$\Gamma_{\mathrm{GC}}(\boldsymbol{k}, z) = \sum_{N=0}^{\infty} z^N \, \Gamma_N(\boldsymbol{k}). \tag{11.16}$$

We obtain

$$\Gamma_{\mathrm{GC}}(\boldsymbol{k}, z) \simeq \left[1 - \frac{z\mathcal{S}}{(2\pi)^d} \, \mathrm{e}^{-k^2 a_0^2/2} \right]^{-1} \propto \frac{1}{z_{\mathrm{c}} - z + ck^2}, \tag{11.17}$$

where we have introduced $c = a_0^2/2$, assumed that $ck^2 \ll 1$, and defined

$$z_{\mathrm{c}} = \frac{(2\pi)^d}{\mathcal{S}}. \tag{11.18}$$

The expression (11.17) is analogous to the Fourier transform of the correlation function of a system close to the critical point, in the Ornstein-Zernike approximation. The fugacity z plays the role of $1/k_B T$, and z_c that of its critical value. When z increases, getting closer to z_c, the average number of monomers in the linear polymer diverges. The analog of susceptibility is given by $\Gamma_{GC}(\boldsymbol{k}{=}0, z)$, which is proportional to the total number of configurations of a linear polymer with fugacity z.

Exercise 11.2 (Polymer on a lattice) Suppose that a polymer is placed over a simple cubic lattice in d dimensions with lattice constant a_0, with the constraint that monomers following along the chain must sit on nearest-neighbor lattice points, but that there is no bound on the number of monomers that can occupy a single lattice point. Show that one obtains an expression analogous to eq. (11.17) for $\Gamma_{GC}(\boldsymbol{k}, z)$ in this case, and calculate the corresponding values of z_c and of the constant c.

Exercise 11.3 Show that if $\Gamma_{GC}(\boldsymbol{k}{=}0, z) \propto |z_c - z|^{-\gamma}$, one has

$$\Gamma_N = \int \mathrm{d}\boldsymbol{r}\, \Gamma_N(\boldsymbol{r}) \propto z_c^{-N} N^{\gamma-1}.$$

In our case, therefore, $\Gamma_N \propto z_c^{-N}$, while the factor $N^{\gamma-1}$ (which is called the **enhancement factor**) is equal to 1, and γ has the classical value 1.

From these results, we can see that the radius of gyration R_G is analogous to the coherence length in critical phenomena. We thus have $R_G \sim |z_c - z|^{-\nu}$, where, in the Gaussian model, $\nu = 1/2$. This corresponds to $R_G \sim N^{\nu}$.

A quantity analogous to the free energy can be obtained by considering the number $\Gamma_N(\boldsymbol{r}{=}0)$ of configurations that return to the origin. This is obviously given by

$$\Gamma_N(\boldsymbol{r}{=}0) = \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^d}\, \Gamma_N(\boldsymbol{k}). \tag{11.19}$$

Exercise 11.4 Show that, for $2 < d < 4$, one has

$$\Gamma(\boldsymbol{r}{=}0, z) = \int \frac{\mathrm{d}\boldsymbol{k}}{(2\pi)^d}\, \Gamma_{GC}(\boldsymbol{k}) \simeq \text{const.} + \text{const.} \times |z_c - z|^{(d-2)/2}.$$

Show that this result implies that for $2 < d < 4$, one has $\Gamma_N(\boldsymbol{r}{=}0) \sim N^{-d/2}$. What occurs for $d = 4$? And what about $d > 4$?

Since $R_G \sim N^{\nu}$ is the analog of ξ, this result implies $2 - \alpha = d\nu = d/2$. Thus we have $\alpha = (d-2)/2$, which is a nonclassical value, but matches that of the Gaussian n-vector model.

11.2.2 The Flory Argument

The results discussed thus far have been obtained by neglecting the excluded volume interactions. One can wonder how important they actually are. A first estimate can be obtained by calculating the number of encounters in a Gaussian polymer—in other words, the

number of times in which monomers that are distant from each other along the chain come to be close to each other in space. We can estimate this quantity by supposing that the polymer is uniformly distributed within a region of space of radius proportional to R_G. The monomer density ρ is therefore proportional to $N/R_G^d \sim N^{1-d/2}$. The encounter probability per unit of volume is proportional to ρ^2, and multiplying it by the volume $\sim R_G^d$, we obtain the number of encounters:

$$\mathcal{N} \sim R_G^d \rho^2 \sim \frac{N^2}{R_G^d} \sim N^{2-d/2}. \tag{11.20}$$

For $d > 4$ we obtain $\lim_{N\to\infty} \mathcal{N} = 0$ (!). This obviously means that the number of encounters remains bounded when the polymer length diverges. On the other hand, if $d < 4$, the number of encounters diverges, and the excluded volume interaction becomes relevant. (For $d = 4$, a more careful analysis shows that the number of encounters increases logarithmically with N.)

A slight modification in this line of reasoning allows us to obtain an excellent estimate of the exponent ν in the presence of excluded volume interactions. Let us consider R_G as a variational parameter and look for the minimum of the free energy of the polymer. The free energy is given by $F = E - TS$, where E is the interaction energy and S the entropy, and we keep the temperature T constant. From the calculation we just performed, we obtain an estimate of E:

$$E \sim \frac{N^2}{R_G^d}. \tag{11.21}$$

The entropy can be estimated by supposing that the polymer is still in a Gaussian ensemble, and as a result, the number of configurations of a width equal to R_G is proportional to $\exp\left(-\text{const.} \times R_G^2/N\right)$. Entropy is the logarithm of this quantity. By taking all the constants as understood, the free energy is therefore given by

$$F = \frac{N^2}{R_G^d} + \frac{R_G^2}{N}. \tag{11.22}$$

For $d > 4$ the first term is negligible compared to the second one, and the effect of the excluded volume interaction is therefore negligible. For $d < 4$, the minimum of $F(R_G)$ is given by

$$R_G \sim N^{3/(d+2)}. \tag{11.23}$$

In this way we obtain the Flory estimate of the exponent ν [Flor53, sec. XIV-1]:

$$\nu = \frac{3}{d+2}, \qquad 1 \le d \le 4. \tag{11.24}$$

This expression is remarkable because it is exact for $d = 1$ ($\nu = 1$) and $d = 2$ ($\nu = 3/4$, which can be obtained with other methods), and it correctly predicts $\nu = 1/2$ for $d = 4$. In $d = 3$, it yields $\nu = 0.6$, to compare with the experimental value $\nu = 0.56 \pm 0.01$. It is not clear, however, why these results are so good. We should add that, up to now, there is no argument of comparable simplicity for the exponent γ that appears in the enhancement factor.

11.2.3 Linear Polymers and Critical Phenomena

We can thus conclude that close analogy exists between the (essentially geometrical) problem of the conformation of a linear polymer with excluded volume interactions and ordinary critical phenomena, discussed in chapter 5. This analogy can be represented by the following mapping:

$$\begin{aligned} N &\longleftrightarrow |T - T_c|^{-1}, \\ R_G &\longleftrightarrow \xi, \\ \Gamma_N(\boldsymbol{r}) &\longleftrightarrow C(\boldsymbol{r}, T), \\ \Gamma_N(\boldsymbol{r}{=}0) &\longleftrightarrow \Delta F. \end{aligned}$$

It can also be put on a more formally explicit foundation. Actually, as de Gennes showed in 1972 (see [Genn79]), the properties of linear polymers with excluded volume interactions correspond to the critical properties of a model with an n-dimensional order parameter, in the limit $n \to 0$.

To derive this relation, we use an argument by G. Sarma [Sarm79]. We start by proving the following geometrical lemma:

Let $\boldsymbol{s} = (s_\alpha)$ be a classical spin with n components and with norm $s^2 = \sum_{\alpha=1}^{n} s_\alpha^2 = n$. Let us consider the moments $\left\langle s_{\alpha_1}^{p_1} \cdots s_{\alpha_n}^{p_n} \right\rangle$ of the components of s, where $\langle \ldots \rangle$ is the average over the sphere of radius $\sqrt{n}$ in n dimensions. Then these moments all vanish in the limit $n \to 0$, except the trivial one (for which $p_i = 0, \forall i$) and those of the form $\langle \sigma_\alpha^2 \rangle$, which are all equal to 1.

This result can be easily obtained by considering the generating function

$$f(\boldsymbol{k}) = \left\langle \exp\left(\mathrm{i} \sum_\alpha k_\alpha s_\alpha \right) \right\rangle. \tag{11.25}$$

This quantity can depend only on k^2, and satisfies the differential equation

$$-\sum_{\alpha=1}^{n} \frac{\partial^2 f}{\partial k_\alpha^2} = \left(\sum_\alpha s_\alpha^2 \right) f = nf. \tag{11.26}$$

In spherical coordinates this equation reads

$$-\left(\frac{\mathrm{d}^2}{\mathrm{d}k^2} + \frac{n-1}{k} \frac{\mathrm{d}}{\mathrm{d}k} \right) f = nf. \tag{11.27}$$

The boundary conditions for this equation are easily obtained:

$$f(0) = 1, \qquad \left. \frac{\partial^2 f}{\partial k_\alpha^2} \right|_{k=0} = -\langle s_\alpha^2 \rangle = -1. \tag{11.28}$$

Therefore one has, for $k^2 \ll 1$,

$$f(\boldsymbol{k}) \simeq 1 - \frac{1}{2} k^2 + \cdots. \tag{11.29}$$

Now, $f(\boldsymbol{k}) = 1 - k^2/2$ is the solution of eq. (11.27) for $n = 0$ and satisfies the initial conditions—it follows that it is the generating function for $n = 0$. By taking the derivative of this generating function with respect to k_α at $k_\alpha = 0$, $\forall\alpha$, we obtain the moments of s_α. We can thus see that the only nonvanishing such moment, apart from the trivial one, is $\langle s_\alpha^2 \rangle = 1$. This concludes the proof.

We now consider a "magnetic" model defined on a simple cubic lattice in d dimensions, where to each lattice point i is associated a vector variable $\boldsymbol{s}_i$ with n dimensions, of magnitude square $\sum_{\alpha=1}^{n} s_\alpha^2 = n$. The weight of a configuration $s = (\boldsymbol{s}_i)$ is given by the Boltzmann-Gibbs factor $\mathrm{e}^{-\mathcal{H}(s)}/Z$, where the Hamiltonian $\mathcal{H}(s)$ is given by

$$\mathcal{H}(s) = -K \sum_{<ij>} (\boldsymbol{s}_i \cdot \boldsymbol{s}_j) , \tag{11.30}$$

in which $K > 0$ and the sum is extended to all nearest-neighbor pairs of sites in the lattice. The partition function Z has the following expression:

$$Z = \int \prod_i \mathrm{d}\boldsymbol{s}_i \, \mathrm{e}^{-\mathcal{H}(s)}, \tag{11.31}$$

where the integrals are extended over the n-dimensional sphere of radius $\sqrt{n}$.

We can expand Z in the following form:

$$Z = \left\langle \prod_{<ij>,\alpha} \left(1 + K s_{i\alpha} s_{j\alpha} + \frac{1}{2} K^2 s_{i\alpha}^2 s_{j\alpha}^2 + \cdots \right) \right\rangle . \tag{11.32}$$

We shall now see that we do not need to expand further. The different terms that contribute to Z in an expansion in powers of K are represented by graphs. A typical term is a *cycle* on the lattice corresponding to a product of $K s_{i\alpha} s_{j\alpha}$, with α fixed. Moreover, because of the geometrical lemma, all the averages except $\langle s_{i\alpha}^2 \rangle$ vanish, and therefore one cannot pass through each point on the lattice more than once: thus the limit $n \to 0$ forces the excluded-volume constraint. Diagrams cannot intersect themselves or be overlapped by other diagrams (see figure 11.1, left). Each cycle contributes K^N, where N is the number of bonds. Summing over α yields nK^N, which tends to zero as n tends to zero. One thus obtains $\lim_{n\to 0} Z = 1$, but it is also clear that $\lim_{n\to 0}(Z-1)/n$ is equal to the number of configurations of a loop polymer over a lattice with excluded volume. Notice that this limit is also given by $\lim_{n\to 0} \ln Z/n$.

Let us now consider the correlation function between, say, $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$, respectively placed at $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$. The correlation function for component 1 has the following expression:

$$C_{11}(\boldsymbol{r}_i, \boldsymbol{r}_j) = \langle s_{i,1} s_{j,1} \, \mathrm{e}^{-\mathcal{H}(s)} \rangle = \left\langle s_{i,1} s_{j,1} \prod_{<\ell m>,\alpha} (1 + K s_{\ell\alpha} s_{m\alpha}) \right\rangle . \tag{11.33}$$

The various contributions to $C_{11}(\boldsymbol{r}_i, \boldsymbol{r}_j)$ are represented by graphs. The contribution of any cycle vanishes when one sums over α, and therefore, only those walks without intersection that connect $\boldsymbol{r}_i$ to $\boldsymbol{r}_j$ remain (see figure 11.1, right). For each bond $<\ell m>$, one has to consider the term $s_{\ell\alpha} s_{m\alpha}$, in which one does not sum over α. Each graph of this type

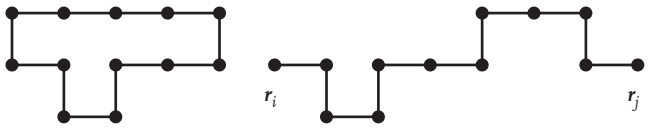


FIGURE 11.1. Graphical expansion in the $n = 0$ model. Left: A typical term contributing to Z in eq. (11.32). Right: A typical term contributing to $C(\boldsymbol{r}_i, \boldsymbol{r}_j)$.

represents the configuration of an excluded volume polymer over the lattice. Each configuration is counted once and only once, and its weight is proportional to K^N, where N is the number of lattice steps that appear in the configuration. Consequently,

$$C_{11}(\boldsymbol{r}_i, \boldsymbol{r}_j) = \left\langle s_{i1} s_{j1} \, \mathrm{e}^{-\mathcal{H}(s)} \right\rangle = \sum_N K^N \, \Gamma_N(\boldsymbol{r}_i - \boldsymbol{r}_j), \tag{11.34}$$

where $\Gamma_N(\boldsymbol{r})$ is the number of excluded volume configurations with N bonds that connect the origin with point $\boldsymbol{r}$ on the lattice. We have thus established a correspondence between the problem of the excluded volume polymer and the zero-component magnetic model. It is easy to obtain the behavior of the various quantities we are interested in in the limit $N \to \infty$ (which corresponds to the critical limit $T \to T_c$ for the magnetic model) from this correspondence.

Let us now consider the magnetic problem in the presence of an external magnetic field $\boldsymbol{h} = (h, 0, \ldots, 0)$. In the limit $n \to 0$, this ensemble describes a solution that contains several excluded volume polymers.

Exercise 11.5 Show that the coefficient of h^{2p} in $\lim_{n\to 0} \ln Z/n$ is given by graphs that represent the configurations of p excluded-volume polymers over the lattice. The contribution of each graph is proportional to K^N, where N is the total number of bonds present in the configuration.

In this manner, h^2 plays the role of fugacity for the polymers, while K plays that of fugacity for the monomers. We note that in this approach, only the total number of monomers is fixed via K, while the length of each polymer is not separately fixed. Therefore, the ensemble described in this fashion contains polymers of all lengths smaller than N. One says that this ensemble describes a **polydisperse** solution.

The physics of polymer solutions at different densities is a fascinating field with many relevant applications. An accessible introduction is due to M. Doi [Doi96]. For a more advanced discussion, I refer the reader to the books by de Gennes [Genn79] and des Cloiseaux and Jannink [Cloi90].

11.3 Percolation

In many situations that deal with random systems, it is necessary to evaluate the connectivity of impurities. A ceramic filter, for example, is composed of a solid matrix in which a certain volume fraction, randomly arranged, is void. In order for the filter to work, the

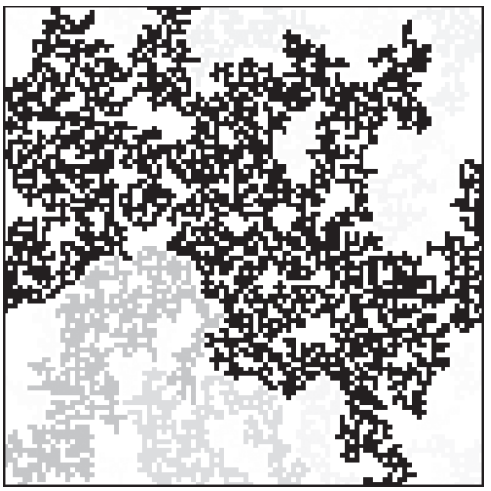


FIGURE 11.2. Site percolation in two dimensions. The sites on a 100×100 square lattice are occupied with a probability $p = 0.59$. Different clusters have different shades of gray, which become darker as their size increases. One observes a large cluster that percolates, connecting the opposite sides of the sample.

voids should connect one side of the filter to the opposite one. In this case one says that the voids **percolate**. One can ask what the minimal volume fraction is of the voids such that this situation arises (with probability close to 1). Analogously, consider a system composed of a random arrangement of conducting and insulating balls. If the fraction of conducting balls is small, the sample will be insulating. If it is large enough, the sample will be conducting. We can ask what the minimal fraction is of conducting balls such that the sample is conducting, and how the conductivity of the sample depends on this fraction. These questions are at the core of **percolation theory**, which was introduced in the 1950s by Hammersley and collaborators [Broa57].

We can formalize the "standard" percolation problem in the following manner. Consider a simple cubic lattice in d dimensions. Let us assume that for each pair of nearest-neighbor lattice sites, a bond is present with probability p, independently for different pairs. We now say that two lattice sites, i and j, are **connected** if there is a path connecting i to j, always going through bonds that are present in the lattice. (For simplicity, we say that each site is connected to itself.) The relation "being connected" is an equivalence relation, since it is reflexive, symmetric, and transitive. The equivalence classes of lattice sites under this relation are called **clusters**. Therefore a cluster, when it is not made of a single isolated site, is made of sets of sites such that it is always possible to go from one site of the cluster to a different one, always jumping from nearest neighbor to nearest neighbor on bonds that are present.

The percolation problem requires us to determine the properties of these clusters as a function of p, and in particular to identify the value of p above which the system exhibits one or more clusters that connect the opposite sides of the sample. This value of p is known as the **percolation threshold**.

What we have just defined is known as the **bond percolation** problem. A closely connected problem is that of **site percolation**, in which the lattice sites can be occupied or empty. In this case, two different sites are connected if there is a path connecting them, always leading from nearest neighbor to nearest neighbor over occupied sites. An example of such a system is shown in fig. 11.2.

There are more general forms of the percolation problem, where, for example, the occupations of different bonds are not independent random variables, or where the connecting paths have a preferred direction. Here we deal with the standard percolation problem, unless mentioned otherwise. For a recent review, see, for example, [Sabe15].

11.3.1 Analogy with Magnetic Phenomena

One can establish an analogy between magnetic phenomena—for example, in the Ising model—and percolation, by supposing that an Ising spin variable is associated with each site, and that if a pair of nearest neighbors site is connected, there is a ferromagnetic interaction of intensity J, where J is very large [Dunn75]. If J is large enough, at the given temperature, all the spins of a given percolation cluster will be oriented in parallel. Let us consider, for example, bond percolation. Then each site belongs to exactly one cluster. When the bond occupation probability p is small, only finite clusters are present, and therefore, in the absence of an external field, the different clusters will be oriented randomly, and the total magnetization of the system vanishes. On the other hand, when p is greater than the percolation threshold p_c, an infinite cluster will be present, and if the interaction J is strong enough, all the spins of this cluster will be oriented in the same direction. Thus, if there is only one infinite cluster, the magnetization per spin will be proportional to the probability $P(p)$ that a site belongs to the infinite cluster. In this manner, one can establish an analogy between $P(p)$ and spontaneous magnetization. Here we assume that the infinite cluster is unique, which can be shown to be true with probability 1 by more sophisticated arguments.

Let us denote by $\nu_s(p)$ the average number of clusters of size s per lattice site. The probability that a specific site belongs to a cluster of size s is given by $s\nu_s(p)$. We can then express $P(p)$ in terms of $\nu_s(p)$ by the relation

$$\sum_s s\nu_s(p) + P(p) = 1. \tag{11.35}$$

Exercise 11.6 Show that the magnetic susceptibility χ per site for this system, with a vanishing magnetic field, is given by

$$\chi = \frac{J}{k_B T} \sum_s s^2 \, \nu_s(p).$$

As a result, χ can be expressed as a function of the average cluster size, defined by

$$S(p) = \frac{\sum_s s^2 \, \nu_s(p)}{\sum_s s \, \nu_s(p)}. \tag{11.36}$$

The partition function Z of the magnetic system is obtained by summing the Boltzmann factor $\exp(-\mathcal{H}/k_B T)$ over all the spin configurations. Let us set the zero of energy so that it vanishes when the spins are parallel, and is much larger than $k_B T$ otherwise. Then the Boltzmann factor is equal to 1 if the connected neighbors point in the same direction, and it practically vanishes otherwise. Therefore, Z is equal to the number of allowed spin

configurations, where a configuration is allowed if all the spins in a connected cluster point in the same direction. The number of allowed configurations is therefore equal to 2^L, where L is the number of clusters. The free energy F is proportional to the logarithm of Z. One therefore obtains

$$F \propto \ln Z \simeq \ln 2 \times L, \tag{11.37}$$

where L is the number of independent clusters that appear in the current bond configuration. Thus the analog of the free energy per site is the average number of clusters per site:

$$C(p) = \frac{\langle L \rangle}{N} = \sum_s \nu_s(p). \tag{11.38}$$

Exercise 11.7 Show that the analog of the magnetic correlation function between site i and site j is given by the probability that i and j belong to the same finite cluster.

11.3.2 Percolation in One Dimension

Here we consider the standard problem of site percolation in one dimension. This problem is slightly pathological, but can be solved exactly.

We have a one-dimensional lattice composed of $N+1$ sites ($i = 0, 1, \ldots, N$). Each site can be present with probability p, independently of the other sites. Let us now consider two sites at distance r from each other. What is the probability $C(r, p)$ that they belong to the same cluster? It is easy to see that they will belong to the same cluster only if they are both occupied and if all the $r-1$ sites that separate them are also occupied. Therefore,

$$C(r, p) = p^{r+1}. \tag{11.39}$$

Let us note that for $r \to \infty$, one always has $C(r, p) \to 0$ except when p is exactly equal to 1. In one dimension, therefore, the probability of having an infinite cluster vanishes for $p < 1$.

Let us now calculate the number $\nu_s(p)$ of clusters of size s per site. To have a cluster of size s there should be s successive occupied sites, flanked by two unoccupied ones. One has therefore

$$\nu_s(p) = p^s(1-p)^2. \tag{11.40}$$

Then the probability that a site belongs to a cluster of size s is given by $s\nu_s(p)$. We can check that $\sum_s s\nu_s(p) = p$. By a simple calculation, and letting the total number of sites N go to infinity, we obtain the average cluster size:

$$S(p) = \frac{\sum_s s^2 \nu_s(p)}{\sum_s s\nu_s(p)} = \frac{1+p}{1-p}. \tag{11.41}$$

Therefore $S(p)$ diverges as $(1-p)^{-1}$.

11.3.3 Percolation on the Bethe Lattice

Percolation in one dimension is disappointing because the infinite cluster appears only when $p = 1$, where the system is no longer random. It is, however, possible to define a

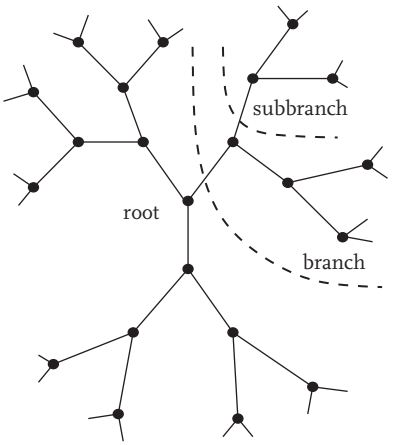


FIGURE 11.3. The neighborhood of the origin in the Bethe lattice. The origin is called the root, and is surrounded by ζ nearest neighbors, which form each the root of a branch. These are connected to the root, and to $\zeta - 1$ nearest neighbors, each forming the root of a subbranch. In the figure, $\zeta = 3$.

problem, almost as simple to solve, for which an infinite cluster arises at a value $p_c < 1$ of the occupation probability. This problem is site percolation on the **Bethe lattice** (also known as the Cayley tree).

The Bethe lattice (see figure 11.3) is an ideal lattice without cycles that is defined recursively. One starts from a site 0, called the root, which is surrounded by ζ nearest neighbors. For each of the nearest neighbors (layer 1 of the lattice), one defines $\zeta - 1$ nearest neighbors (layer 2); and then for each site in layer 2, one defines $\zeta - 1$ nearest neighbors (layer 3). This procedure is iterated without limit. Then layer ℓ contains $\zeta\,(\zeta - 1)^{\ell}$ sites, wheras the total number of sites is $\zeta\,(\zeta - 1)^{\ell+1}$. At each step, therefore, a finite fraction of sites belongs to the last layer. If we consider the points contained inside a sphere or a box in a lattice of finite dimension d, it is easy to convince oneself that the ratio between the points contained on the surface and those contained in the volume goes to zero as the volume goes to infinity. Since this ratio tends to a finite value in the Bethe lattice, one can only realize it in a space of infinite dimensionality. We denote the probability that an occupied site belongs to an infinite cluster by $P(p)$. Let us now show that there exists a value $p_c < 1$ such that, if $p > p_c$, $P(p) > 0$.

Let us consider one of the ζ nearest neighbors of the origin, and denote by $Q(p)$ the probability that it is *not* connected to infinity via its branch. This happens either because the site is not occupied, or because it is occupied, but all its $\zeta - 1$ neighbors on the branch are not connected to infinity via their respective subbranches. Now, these subbranches are independent of one another, and each reproduces, for their root, the same structure. Thus the probability that all these nearest neighbors are not connected to infinity via their subbranches is equal to $Q(p)^{\zeta-1}$. We thus obtain the equation

$$Q = 1 - p + pQ^{\zeta-1}. \tag{11.42}$$

This equation always admits the solution $Q = 1$, but it is easy to see that if one assumes $Q \neq 1$, one obtains the equation

$$\frac{1}{p} = \frac{1 - Q^{\zeta-1}}{1 - Q}. \tag{11.43}$$

This equation admits a solution with $Q < 1$ if $p > (\zeta - 1)^{-1}$. The probability $P(p)$ that the origin, if occupied, belongs to the infinite cluster is equal to the probability that at least one of its ζ nearest neighbors is connected to infinity. We therefore have

$$P(p) = 1 - Q^{\zeta}. \tag{11.44}$$

Obviously, $P(p) > 0$ if $Q(p) < 1$, which occurs, as we saw, when $p > p_c$. We have thus identified the percolation threshold p_c for the Bethe lattice with coordination number ζ:

$$p_c = \frac{1}{\zeta - 1}. \tag{11.45}$$

When p is slightly larger than p_c, $1 - Q(p)$ and, therefore, $P(p)$ are roughly proportional to $p - p_c$. As we have seen, $P(p)$ plays the role of the order parameter of the percolation transition, and p that of the inverse temperature. We can therefore introduce the exponent β by means of the relation

$$P(p) \sim (p - p_c)^{\beta}, \qquad p > p_c. \tag{11.46}$$

We thus obtain that, for the Bethe lattice, one has $\beta = 1$.

It is possible to evaluate the average cluster size $S(p)$ in an analogous way. If we denote the average size of clusters contained in one branch of the Bethe lattice with $T(p)$, it is easy to see that on the one hand, one has for $p < p_c$,

$$S(p) = p\,(1 + \zeta\, T(p))\,, \tag{11.47}$$

while on the other hand,

$$T(p) = p\,[1 + (\zeta - 1)T(p)]\,. \tag{11.48}$$

Exercise 11.8 Prove the previous two relations for $S(p)$ and $T(p)$.

We thus obtain

$$S(p) = \frac{p(1+p)}{1 - (\zeta - 1)p}, \qquad \text{for } p < p_c, \tag{11.49}$$

whereas $S(p) = \infty$ for $p \geq p_c$.

The distribution $\nu_s(p)$ of cluster sizes can be obtained by observing that the probability that a set of s connected Bethe lattice points form a cluster is equal to the probability p^s that they are all occupied, times the probability $(1-p)^{s'}$ that the s' points that make up its boundary are all empty. In the Bethe lattice, s' does not depend on the cluster's shape and is equal to

$$s' = (\zeta - 2)s + 2. \tag{11.50}$$

Exercise 11.9 Prove the previous relation.

To evaluate $\nu_s(p)$, we need to estimate the number of sets of Bethe lattice points that are connected to each other, which is not an easy task. However, if we only want to calculate the ratio between $\nu_s(p)$ and $\nu_s(p_c)$, this number cancels out between numerator and denominator, and we obtain

$$\begin{aligned}\frac{\nu_s(p)}{\nu_s(p_c)} &= \left(\frac{p}{p_c}\right)^s \left(\frac{1-p}{1-p_c}\right)^{(\zeta-2)s+2} \\ &= \left(\frac{p(1-p)^{\zeta-2}}{p_c(1-p_c)^{\zeta-2}}\right)^s \left(\frac{1-p}{1-p_c}\right)^2 .\end{aligned} \tag{11.51}$$

For $p \neq p_c$, this ratio decays exponentially in s: $\nu_s(p)/\nu_s(p_c) \propto \exp[-c(p)s]$, where

$$c(p) = -\ln\left[\frac{p}{p_c}\left(\frac{1-p}{1-p_c}\right)^{\zeta-2}\right]. \tag{11.52}$$

For $p \simeq p_c$, one has $c(p) \sim (p-p_c)^2$. We have thus obtained

$$\nu_s(p) \propto \nu_s(p_c)\, \mathrm{e}^{-c(p)s}, \tag{11.53}$$

where $c(p)$ vanishes for $p \to p_c$ as a power law

$$c(p) \sim \left|p - p_c\right|^{1/\sigma}, \tag{11.54}$$

which defines the **Fisher exponent** σ. This exponent is equal to $1/2$ on the Bethe lattice. Let us suppose that $\nu_s(p_c)$ decays as a power of s:

$$\nu_s(p_c) \sim s^{-\tau}, \qquad s \gg 1. \tag{11.55}$$

This relation defines the **Fisher exponent** τ. We thus obtain

$$\nu_s(p) \sim s^{-\tau}\, \mathrm{e}^{-c(p)s}. \tag{11.56}$$

We can now calculate the average cluster size $S(p)$ for $p < p_c$. Since $\nu_s(p)$ is the average number of clusters of size s per lattice site, the probability that a site belongs to a cluster of size s is given by $s\nu_s(p)$, and each of these clusters contributes s to the average. We thus obtain $S(p) = \sum_s s^2 \nu_s(p)$, while, for $p < p_c$, $\sum_s \nu_s(p) = 1$ since $P(p) = 0$. Therefore

$$S(p) = \sum_s s^2 \nu_s(p) \propto \sum_s s^{-\tau}\, \mathrm{e}^{-c(p)s} \propto [c(p)]^{\tau-3}. \tag{11.57}$$

Thus for p close to, but smaller than, p_c we have $S \sim [c(p)]^{\tau-3} \sim \left|p - p_c\right|^{(\tau-3)/\sigma}$. On the other hand, we had already obtained $\sigma = 1/2$ and $S \sim \left|p - p_c\right|^{-1}$. We thus obtain the value of τ:

$$\tau = \frac{5}{2}. \tag{11.58}$$

For $p > p_c$, one has $\sum_s \nu_s(p) = 1 - P(p)$. Now, if we define $S(p)$ as the average size of the finite clusters only, we obtain $S = \sum_s s^2 \nu_s(p)$ also for $p > p_c$, and therefore $S(p) \sim \left|p - p_c\right|^{-1}$ also in this case.

These relations can be summarized in a scaling scheme. The order parameter is P, and the exponent β is defined by the relation $P \sim |p_c - p|^{\beta}$ for $p \to p_c^+$. The average cluster size S is the analog of the susceptibility, and we can therefore set $S \sim |p_c - p|^{-\gamma}$. Introducing the exponents σ and τ from the cluster distribution $\nu_s(p)$ (equations (11.54) and (11.55)), we obtain the relations

$$\beta = \frac{\tau - 2}{\sigma}, \qquad \gamma = \frac{3 - \tau}{\sigma}. \tag{11.59}$$

We thus have $\beta = 1$, $\gamma = 1$ for the Bethe lattice. More generally, we can consider all the moments M_k of the cluster distribution:

$$M_k(p) = \sum_{s=1}^{\infty} s^k \, \nu_s(p). \tag{11.60}$$

Exercise 11.10 Show that, if $k > \tau - 1$, one has $M_k(p) \sim |p - p_c|^{\Delta_k}$, where $\Delta_k = (\tau - 1 - k)/\sigma$.

If the sum in eq. (11.60) does not diverge for $p \to p_c$, it contains both a part analytic in $p - p_c$ and a singular part that behaves like a power law, in analogy with thermodynamic observables in the neighborhood of the critical point. Let us consider, for instance, the case $k = 0$—in other words, the average number of clusters per site, which is the analog of the free energy density. Considering M_0 as a function of c, we have

$$\frac{\mathrm{d}^2 M_0}{\mathrm{d}c^2} = M_2 \sim c^{\tau - 1}. \tag{11.61}$$

We thus obtain

$$M_0 = M_0(p_c) + \text{const.} \times c(p) + \text{const.} \times c^2(p) + \text{const.} \times c(p)^{\tau - 1}. \tag{11.62}$$

The first three terms are the regular part of M_0. The singular part of M_0 therefore behaves as $c^{\tau - 1} \sim |p - p_c|^{(\tau - 1)/\sigma}$. This exponent is usually defined as $2 - \alpha$. We thus obtain

$$2 - \alpha = \frac{\tau - 1}{\sigma}, \tag{11.63}$$

which is equal to 3 for the Bethe lattice, so that one has $\alpha = -1$, a value that satisfies the scaling law $\alpha + 2\beta + \gamma = 2$. If we consider these as the classical exponents, the Toulouse criterion (see sec. 6.10) yields, via the scaling law $2 - \alpha = d\nu$, the upper critical dimension $d_c = 6$.

11.3.4 Relation with the Potts Model

Just as a relation exists between the problem of polymers in a solution and the critical behavior of a magnetic system, in the limit $n \to 0$, it is also possible to relate the percolation problem to an analogous limit in a specific statistical system: the Potts model. This relation allows one to establish on a more rigorous foundation the analogy between percolation

and critical phenomena, but on the other hand, it is not very useful for the purposes of calculation. I discuss these results by following Lubensky's exposition [Lube79].

The Potts model is a generalization of the Ising model. A variable σ_i, which can assume one of q values $\sigma_i \in \{1, 2, \ldots, q\}$, is assigned to each site of, say, a simple cubic lattice in d dimensions. Each time the σ_i pertaining to nearest-neighbor sites have the same value, the energy is decreased by a certain amount. We also add an external field $h_0 > 0$, which favors the state $\sigma_i = 1$. It is convenient to define the Hamiltonian for the configuration $\sigma = (\sigma_i)$ in the following fashion:

$$\mathcal{H}(\sigma) = -\sum_{<ij>} J \left(q\,\delta_{\sigma_i\sigma_j} - 1\right) - \sum_i h_0 \left(q\,\delta_{\sigma_i 1} - 1\right), \tag{11.64}$$

where the first sum runs over all the nearest-neighbor pairs.

We can now evaluate the partition function Z_N by summing over the different configurations of a lattice containing N sites and $N_\mathrm{b} = \zeta N/2$ bonds, where ζ is the coordination number of the lattice. It is useful to introduce a factor q^{-N} into the definition of Z_N. Defining $K = J/k_\mathrm{B}T$, $h = h_0/k_\mathrm{B}T$ and $p = 1 - \mathrm{e}^{-qK}$, we obtain

$$Z_N = q^{-N} \sum_\sigma \exp\left(-\mathcal{H}(\sigma)/k_\mathrm{B}T\right) = \exp\left[(q-1)(KN_\mathrm{b} + hN)\right] Z'_N, \tag{11.65}$$

where

$$\begin{aligned} Z'_N &= q^{-N} \sum_\sigma \exp\left[qK \sum_{<ij>} \left(\delta_{\sigma_i\sigma_j} - 1\right) + qh \sum_i \left(\delta_{\sigma_i 1} - 1\right)\right] \\ &= q^{-N}(1-p)^{N_\mathrm{b}} \sum_\sigma \prod_{<ij>} \left(1 + \frac{p}{1-p}\delta_{\sigma_i\sigma_j}\right) \exp\left[qh \sum_i \left(\delta_{\sigma_i 1} - 1\right)\right]. \end{aligned} \tag{11.66}$$

Each term in the expansion (11.66) is represented by a graph $\mathcal{G}$ on the lattice, such that the presence of a bond yields a factor $(p/(1-p))\,\delta_{\sigma_i\sigma_j}$ and its absence a factor 1. In this way the lattice splits into connected clusters, and the Potts variables in each cluster must have the same value because of the Kronecker delta factors. Thus the contribution of a cluster with s sites and b bonds, after the sum over the σ_i, is equal to

$$\left(\frac{p}{1-p}\right)^b \left[1 + (q-1)\,\mathrm{e}^{-qhs}\right].$$

Isolated sites are counted as clusters that contain a single site.

We will now denote the number of clusters of the graph $\mathcal{G}$ that have s sites and b bonds by $N\mathcal{K}(\mathcal{G}, b, s)$. We then have

$$\begin{aligned} Z'_N = q^{-N}(1-p)^{N_\mathrm{b}} \sum_\mathcal{G} \Bigg\{ &\left(\frac{p}{1-p}\right)^{N\sum_{s,b} b\,\mathcal{K}(\mathcal{G},b,s)} \\ &\times \prod_{b,s} \left[1 + (q-1)\mathrm{e}^{-qhs}\right]^{N\mathcal{K}(\mathcal{G},b,s)} \Bigg\}. \end{aligned} \tag{11.67}$$

The total number of clusters of graph $\mathcal{G}$ is given by

$$N_c(\mathcal{G}) = \sum_{b,s} \mathcal{K}(b,s), \tag{11.68}$$

and the total number of occupied bonds in $\mathcal{G}$ has the expression

$$\mathcal{B}(\mathcal{G}) = \sum_{b,s} b\mathcal{K}(b,s). \tag{11.69}$$

We then have

$$\begin{aligned} Z'_N &= q^{-N} \sum_{\mathcal{G}} p^{\mathcal{B}(\mathcal{G})} (1-p)^{N_b - \mathcal{B}(\mathcal{G})} \prod_{b,s} \left[1 + (q-1)e^{-qhs}\right]^{N\mathcal{K}(\mathcal{G},b,s)} \\ &= q^{-N} \sum_{\mathcal{G}} P(\mathcal{G}) \prod_{b,s} \left[1 + (q-1)e^{-qhs}\right]^{N\mathcal{K}(\mathcal{G},b,s)}, \end{aligned} \tag{11.70}$$

where $P(\mathcal{G}) = p^{\mathcal{B}(\mathcal{G})} (1-p)^{N_b - \mathcal{B}(\mathcal{G})}$ is the probability of occurrence of the graph $\mathcal{G}$ when the occupation probability of a bond is equal to p.

This fundamental result was derived in 1972 by Fortuin and Kasteleyn [Fort72]. We can exploit it to express the characteristic quantities of percolation in terms of Z_N.

Limit $q \to 1$ and Percolation

It is obvious from eq. (11.70) that in the $q \to 1$ limit, $Z_N = 1$, and therefore $\ln Z_N = 0$. Let us consider the following expression of the free energy per site:

$$f = \lim_{N\to\infty} \lim_{q\to 1} \frac{1}{N(q-1)} \ln Z_N = \lim_{N\to\infty} \left. \frac{\partial \ln Z_N}{\partial q} \right|_{q=1}. \tag{11.71}$$

From equations (11.65) and (11.70), we obtain

$$\begin{aligned} f &= -1 + \frac{\zeta}{2} K + h + \sum_{\mathcal{G}} P(\mathcal{G}) \sum_{b,s} \mathcal{K}(\mathcal{G},b,s)\, e^{-hs} \\ &= -1 + \frac{\zeta}{2} K + h + \sum_{s} \mathcal{K}(s)\, e^{-hs}, \end{aligned} \tag{11.72}$$

where

$$\mathcal{K}(s) = \sum_{\mathcal{G}} P(\mathcal{G}) \sum_{b} \mathcal{K}(\mathcal{G},b,s) \tag{11.73}$$

is the average number per site of clusters of size s. One should note that only clusters containing a finite number of sites contribute to eq. (11.72) in the limit $N \to \infty$, so that all the sums are restricted to finite values of s and b.

In the limit $h \to 0$ we obtain

$$f = -1 - \frac{\zeta}{2} \ln(1-p) + \frac{\langle N_c \rangle}{N}, \tag{11.74}$$

where $\langle N_c \rangle = N \sum_s \mathcal{K}(s)$ is the average number of clusters per site.

The first derivative of f with respect to h gives us probability $P(p)$ that a site belongs to the infinite cluster,

$$\lim_{h\to 0}\frac{\partial f}{\partial h}=\lim_{h\to 0}\left(1-\sum_s s\mathcal{K}(s)\mathrm{e}^{-hs}\right)=1-\sum_s s\mathcal{K}(s)=P(p), \tag{11.75}$$

since $s\mathcal{K}(s)$ is the probability that a site belongs to a cluster of size s and the sum runs over only finite clusters.

The second derivative of f gives us the average cluster size $S(p)$:

$$\left.\frac{\partial^2 f}{\partial h^2}\right|_{h=0}=\sum_s s^2\mathcal{K}(s). \tag{11.76}$$

Let us now evaluate the correlation function of the Potts model between two sites, respectively located at $\boldsymbol{r}$ and $\boldsymbol{r}'$:

$$(q-1)G(\boldsymbol{r},\boldsymbol{r}')=\left\langle q\delta_{\boldsymbol{r}\boldsymbol{r}'}-1\right\rangle, \tag{11.77}$$

where the factor $(q-1)$ has been introduced for convenience. Equation (11.70) reduces, in the $h\to 0$ limit, to

$$Z'_N=\sum_{\mathcal{G}}P(\mathcal{G})q^{N_c(\mathcal{G})-N}. \tag{11.78}$$

Given two sites, placed at $\boldsymbol{r}$ and $\boldsymbol{r}'$, we can separate the graphs into two sets: the first $\mathcal{G}_1$, in which $\boldsymbol{r}$ and $\boldsymbol{r}'$ belong to the same cluster, and the second $\mathcal{G}_2$, in which they belong to different clusters. If $\mathcal{G}\in\mathcal{G}_1$, multiplying by $\delta_{\sigma_{\boldsymbol{r}_1}\sigma_{\boldsymbol{r}_2}}$ does not change anything, while if $\mathcal{G}\in\mathcal{G}_2$, this multiplication decreases the number of clusters by 1. Therefore,

$$\begin{aligned}(q-1)G(\boldsymbol{r},\boldsymbol{r}')&=\left\langle q\delta_{\boldsymbol{r}\boldsymbol{r}'}-1\right\rangle\\&=\frac{1}{Z'_N}\left[\sum_{\mathcal{G}\in\mathcal{G}_1}(q-1)P(\mathcal{G})\,q^{N_c(\mathcal{G})-N}+\sum_{\mathcal{G}\in\mathcal{G}_2}(q^{-1}q-1)P(\mathcal{G})\,q^{N_c(\mathcal{G})-N}\right]\\&=\frac{q-1}{Z'_N}\sum_{\mathcal{G}_1}P(\mathcal{G})\,q^{N_c(\mathcal{G})-N}\\&=\frac{q-1}{Z'_N}\sum_{\mathcal{G}}P(\mathcal{G})\,q^{N_c(\mathcal{G})-N}\Delta(\boldsymbol{r},\boldsymbol{r}')=(q-1)\left\langle\Delta(\boldsymbol{r},\boldsymbol{r}')\right\rangle,\end{aligned} \tag{11.79}$$

where $\Delta(\boldsymbol{r},\boldsymbol{r}')$ is equal to 1 if the sites in $\boldsymbol{r}$ and $\boldsymbol{r}'$ belong to the same cluster, and to zero otherwise. We thus obtain

$$\lim_{q\to 1}G_q(\boldsymbol{r},\boldsymbol{r}')=\sum_{\mathcal{G}}P(\mathcal{G})\,\Delta(\boldsymbol{r},\boldsymbol{r}')=\left\langle\Delta(\boldsymbol{r},\boldsymbol{r}')\right\rangle. \tag{11.80}$$

The average probability that the sites placed in $\boldsymbol{r}$ and $\boldsymbol{r}'$ belong to the same *finite* cluster is given by

$$C(\boldsymbol{r},\boldsymbol{r}')=\left\langle\Delta(\boldsymbol{r},\boldsymbol{r}')\right\rangle-P^2(p). \tag{11.81}$$

Thus the percolation transition is described by the $q \to 1$ limit of the Potts model, and the analogy with critical phenomena is established in this way in a solid footing, with $P(p)$ playing the role of the order parameter, and $C(\boldsymbol{r}, \boldsymbol{r}')$ that of the correlation function.

Limit $q \to 0$ and Tree Statistics

In the limit $q \to 0$, we exploit the fundamental relation, valid for any graph $\mathcal{G}$,

$$\mathcal{B}(\mathcal{G}) + N_{\mathrm{c}}(\mathcal{G}) - N_{\ell}(\mathcal{G}) = N, \tag{11.82}$$

where N_{c} is the number of clusters and $N_{\ell}(\mathcal{G})$ is the number of closed loops. We also have $p = qK + \mathrm{O}(q^2)$. We thus obtain the following expression for Z'_N in the $q \to 0$ limit:

$$\lim_{q\to 0} Z'_N = \lim_{q\to 0} \sum_{\mathcal{G}} q^{N_{\ell}(\mathcal{G})} \prod_{b,s} (1+hs)^{N\mathcal{K}(\mathcal{G},b,s)} = \sum_{\mathcal{T}} \prod_{b,s} (1+hs)^{N\mathcal{K}(\mathcal{T},b,s)} = \mathcal{Z}'_N(h). \tag{11.83}$$

Here $\mathcal{T}$ runs on the set of **trees** on the lattice, namely, graphs without loops, and $N\mathcal{K}(\mathcal{T}, b, s)$ is the number of clusters of size s with b bonds in $\mathcal{T}$.

The free energy per site is given by

$$f_0 = \lim_{N\to\infty} \lim_{q\to 0} \frac{\ln Z}{N(q-1)} = \frac{\zeta}{2} K + k - \lim_{N\to\infty} \mathcal{Z}'_N. \tag{11.84}$$

The first and second derivatives of f_0 with respect to h are respectively given by

$$\frac{\partial f_0}{\partial h} = 1 - \frac{1}{\mathcal{Z}'_N} \sum_{\mathcal{T}} K^{\mathcal{B}(\mathcal{T})} \sum_{s} s\mathcal{K}(\mathcal{T}, s) = 1 - \langle s \rangle_{\mathrm{t}}, \tag{11.85}$$

$$\frac{\partial^2 f_0}{\partial h^2} = \frac{1}{\mathcal{Z}'_N} \sum_{\mathcal{T}} K^{\mathcal{B}(\mathcal{T})} \sum_{s} s^2\mathcal{K}(\mathcal{T}, s) = \left\langle s^2 \right\rangle_{\mathrm{t}}, \tag{11.86}$$

where $\langle s \rangle_{\mathrm{t}}$ and $\left\langle s^2 \right\rangle_{\mathrm{t}}$ are the average and the second moment of the tree sizes, evaluated with respect to the distribution $K^{\mathcal{B}(\mathcal{T})}/\mathcal{Z}'_N$.

Exercise 11.11 Show that $\lim_{q\to 0} G(\boldsymbol{r}, \boldsymbol{r}') = \left\langle \Delta(\boldsymbol{r}, \boldsymbol{r}') \right\rangle_{\mathrm{t}}$, where $\langle \ldots \rangle_{\mathrm{t}}$ represents the average with respect to the distribution $K^{\mathcal{B}(\mathcal{T})}/\mathcal{Z}'_N$. Moreover, show that in the limit $K \to \infty$, one obtains the average with respect to the **spanning trees**—in other words, to those trees that visit all the sites.

It is possible to map the Potts model into a suitable n-dimensional vector model, where $n = q - 1$, with a different symmetry than for ordinary critical phenomena. This model represents a starting point for the expansion of the critical exponents in $d = 6 - \epsilon$ dimensions, which is, however, not very useful from the point of view of actual calculation. For this purpose, the fixed dimension renormalization schemes of percolation, which we will discuss shortly, are more useful.

The Fortuin-Kasteleyn representation (11.70) has been exploited to define a clever Monte Carlo algorithm for Ising systems, derived by Swendsen and Wang [Swen87] and Wolff [Wolf89], following a result by Coniglio and Klein [Coni80]. Consider a spin

configuration $\sigma = (\sigma_i)$ of an Ising model in d dimensions. For each pair $<ij>$ of nearest neighbors, introduce a bond with probability $p = e^{-2J/k_B T}$ (where J is the coupling constant) if $\sigma_i = \sigma_j$, and with probability 0 otherwise. We thus split the lattice into a set of connected clusters. Then the Metropolis algorithm (see sec. 8.4.2) is applied to moves in which all the spins in each cluster are flipped in one go. Once the new configuration is obtained, a new bond distribution is assigned. One can show that the resulting algorithm satisfies detailed balance with respect to the Ising Hamiltonian, and leads therefore to the correct configuration sampling. In this way one obtains moves that change spins in a coordinated way over large spans of the lattice, decreasing the correlation between successive configurations of the system. Thus the approach to equilibrium is faster and the statistics for the evaluation of observables are improved (see, for example, [Krau06, sec. 5.2.3]).

11.3.5 Renormalization in Real Space

Let us consider bond percolation over a square lattice in two dimensions. We want to renormalize the bonds on an elementary 2×2 cell (see figure 11.4). The renormalization scheme is presented in the illustration. It can be interpreted supposing that the renormalized AC bond is present only if one can go from A or D to C or F in the original system. Analogously, the renormalized AG bond is considered present if one can go from A or B to G or H. Let us now consider the renormalized AC bond—the AB, DE, BE, BC, and EF bonds contribute to it in the original system. The renormalized AC bond will be present in the configurations shown in the illustration, to which the relative probabilities are associated. One thus obtains the recurrence relation that corresponds to a rescaling by a factor $b = 2$:

$$\begin{aligned} p' &= p^5 + 5p^4(1-p) + 8p^3(1-p)^2 + 2p^2(1-p)^3 \\ &= 2p^5 - 5p^4 + 2p^3 + 2p^2. \end{aligned} \tag{11.87}$$

In addition to the two trivial fixed points $p = 0$ and $p = 1$, this relation admits the fixed point $p = p^* = 1/2$, which is consistent with the exact value of p_c for the square lattice. One additionally has

$$\lambda = \left.\frac{dp'}{dp}\right|_{p^*} = \frac{13}{8}, \tag{11.88}$$

from which it follows that the fixed point p^* is unstable, and that

$$\nu = \frac{\ln b}{\ln \lambda} \simeq 1.4277, \tag{11.89}$$

to compare with the value $4/3$, which is presumed to be exact in two dimensions.

This is not the only method one can use to set up the renormalization program in real space. Indeed, one can obtain a fairly simple recurrence relation by modifying the lattice structure in a suitable fashion before renormalization, so as to make the renormalization itself as close to a one-dimensional transformation as possible. Let us focus our attention on the origin in a d-dimensional cubic lattice, and assume that the scaling factor is b.

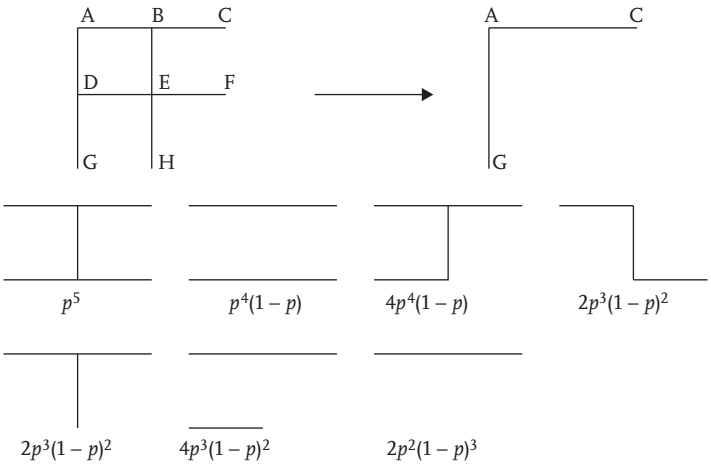


FIGURE 11.4. The configurations that contribute to the renormalization of bond percolation in the case of a 2×2 cell.

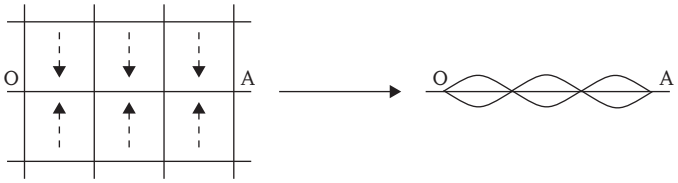


FIGURE 11.5. Migdal-Kadanoff transformation. Scheme showing the shift in renormalization bonds for $b = 3$ and $d = 2$.

So the first vertex that remains after renormalization, in direction x, is located b lattice steps away from the origin. We denote this vertex by A. Let us shift the bonds parallel to the x axis that connect vertices in the $(d-1)$ transversal directions, in such a manner that they connect vertices sequential to the origin along the x axis. In this way, the connection between the origin and A is composed of a sequence of b "bundles" of bonds, each one composed of $1+(d-1)(b-1)$ bonds. This process is shown schematically in figure 11.5. This type of recurrence relations was introduced by A. A. Migdal [Migd75] and generalized by Leo P. Kadanoff [Kada76], and is therefore known as the Migdal-Kadanoff transformation.

We can therefore apply the recurrence relation in a direct fashion: the bond between the origin and A in the renormalized system is present if, for each of the b bundles, at least one other bond is present. This leads to the following recurrence relation:

$$p \longrightarrow p' = \left[1-(1-p)^{[1+(d-1)(b-1)]}\right]^b. \tag{11.90}$$

It is better to take the $b \to 1$ limit and calculate $\partial p'/\partial b$:

$$W(p) = \left.\frac{\partial p'}{\partial b}\right|_{b=1} = p\ln p - (d-1)(1-p)\ln(1-p). \tag{11.91}$$

For $d = 2$, we obtain $W(p) = p \ln p - (1-p)\ln(1-p)$, from which one sees that there is an unstable fixed point for $p = p^* = 1/2$. For this point, one has

$$\lambda = \left.\frac{\mathrm{d}W}{\mathrm{d}p}\right|_{p^*} \simeq 0.6137, \tag{11.92}$$

which corresponds to an exponent $\nu = 1/\lambda \simeq 1.6294$, to be compared to the value $\nu = 4/3$, presumed to be exact.

Exercise 11.12 Calculate the exponent ν for $d = 3$, and compare it with the numerical value $\nu \simeq 0.9$.

Exercise 11.13 (Migdal-Kadanoff renormalization of the Ising model) The Migdal-Kadanoff recurrence relation can also be derived for more "traditional" statistical systems. Once the bonds have been shifted, one obtains a one-dimensional system for which decimation (see sec. 6.2) is exact. Apply this method to the d-dimensional Ising model, and obtain the recurrence relation for $K = J/k_B T$ corresponding to a scaling factor b. Calculate the equation for the infinitesimal transformation corresponding to $b \to 1^+$. Identify the fixed point for $d = 2$ and $d = 3$, and compare the value for the critical temperature and the exponent ν obtained in this way with the exact solution or available numerical values. Discuss the limits $d \to 1$ and $d \to 3$.

11.3.6 Fractal Structure of a Percolation Cluster

Percolative systems provide a good example of the relevance of the concept of fractals introduced by B. Mandelbrot [Mand77], which we have already briefly seen in section 9.3. Fractal objects are geometrical objects that contain detailed structure at smaller and smaller scales. They are often endowed with a fractal dimension that can assume non-integer values. A classic example of a fractal object immersed in one dimension is the Cantor set: it is obtained by considering the closed interval $[0, 1]$ and deleting its (open) middle third—namely, the interval $(1/3, 2/3)$. One thus obtains two closed intervals $[0, 1/3]$ and $[2/3, 1]$. The procedure is iterated, removing from these two intervals their middle intervals, and so forth. The points that remain at the end of this construction constitute the Cantor set: they can be characterized as the points between 0 and 1 such that the expansion in base 3 of their coordinate does not contain the digit 1. The resulting set is a dust of points that does not contain any interval. We can define its dimension d_H (called the **Hausdorff dimension**) as follows. Let us cover the interval $[0, 1]$ with intervals of length ℓ. The number $\mathcal{N}(\ell)$ of intervals that contain points belonging to the set grows as a power of ℓ as ℓ decreases: $\mathcal{N}(\ell) \sim \ell^{-d_H}$. In our case, each time the length ℓ is divided by 3, the number $\mathcal{N}(\ell)$ is multiplied by 2. We thus obtain $d_H = \ln 2/\ln 3 = 0.630927\ldots$, a number between 0 and 1, as we anticipated.

Exercise 11.14 Calculate the Hausdorff dimension of the **Sierpiński triangle** (see figure 11.6, left), recursively defined as follows. Consider an equilateral triangle, and divide it into four equal triangles. The middle triangle is removed, and the construction

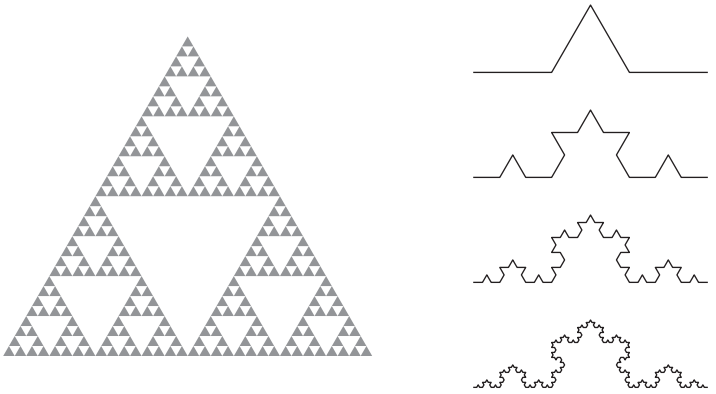


FIGURE 11.6. Examples of fractals. Left: Sierpiński triangle, up to five iterations of the procedure. Right: The first four steps in the construction of the von Koch curve.

is repeated for the remaining three triangles and so forth. Note that the dimension d_{H} obtained in this manner satisfies $1 < d_{\mathrm{H}} < 2$.

Exercise 11.15 Calculate the Hausdorff dimension of the **von Koch curve** (see figure 11.6, right), recursively defined as follows. Consider a segment of unit length in the plane. The middle third of the segment is replaced with two segments of length 1/3 each, as in the figure. The construction is repeated for each of the four segments thus obtained.

The introduction of the concept of fractal object has allowed us to quantitatively describe the properties of a great number of natural or mathematical objects that were previously considered too irregular to be treated mathematically. These objects are defined because they possess, at least approximately, a symmetry under scale transformations. For instance, the triangle at the top of the Sierpiński triangle is a 1:2 scale copy of the Sierpiński triangle itself.

An alternative definition of the Hausdorff dimension, which works well in these simple cases, is obtained by evaluating the number of points in the set contained within an ℓ radius from a point belonging to the set. This obviously makes sense if ℓ is sufficiently larger than the point size, set, say, by the lattice constant a_0.

Let us consider, for example, a smooth D-dimensional manifold immersed in d-dimensional space. If we take a point that belongs to this manifold and trace a sphere of radius ℓ with center at that point, the volume of the part of the manifold contained in this sphere will be proportional to ℓ^D and will contain a number of points proportional to ℓ^D/a_0^D. This confirms our intuitive idea that the manifold has D dimensions. Let us now suppose that we are dealing with an approximation to a Cantor set, constructed by iterating the basic construction a finite (but large) number of times, so that the smallest intervals are of length a_0. The number of points contained within a circle of radius ℓ then grows as $\ell^{\ln 2/\ln 3}$, because if ℓ is multiplied by 3, the number of points doubles. We can thus see that this definition of dimension gives us (in this case) the same result as the preceding definition.

Exercise 11.16 Derive the Hausdorff dimensions of the Sierpiński triangle and of the von Koch curve according to this definition.

We saw that the critical fluctuations also possess an analogous property of symmetry under scale transformations. Similarly, the set of occupied points in the lattice in a site percolation problem, at the percolation threshold p_c, is a system that admits a symmetry under scale transformations, at least as long as the scales being considered are large enough when compared to the lattice step. We thus come upon the idea of describing the infinite cluster at the percolation threshold (the so-called incipient percolation cluster) as a fractal characterized by a certain Hausdorff dimension.

The scale invariance is obviously only approximate. On the one hand, it is no longer true if the lengths we are interested in are of the order of the lattice step a_0. On the other hand, if we are not exactly at the percolation threshold, the cluster is no longer fractal over distances of the order of the typical cluster linear size ξ. Between these two lengths, however, there can be some orders of magnitude of distance for which a fractal description is appropriate.

Let us now consider a site percolation model in d dimensions. Let us suppose that p is close to, but slightly larger than, the percolation threshold p_c. We consider a point belonging to the infinite cluster and evaluate the number $\mathcal{M}(\ell)$ of points of the infinite cluster that are contained within a d-dimensional sphere of radius ℓ from the point. In general, we expect $\mathcal{M}(\ell)$ to be described as a scaling expression, as a function of ℓ and the coherence length ξ:

$$\mathcal{M}(\ell) = \ell^{\psi} \widehat{\mathcal{M}}(\ell/\xi). \tag{11.93}$$

If $\ell \gg \xi$ we expect the irregularities of the infinite cluster, on average, to disappear, and therefore that

$$\mathcal{M}(\ell) \sim P\, \ell^{d}, \tag{11.94}$$

where P is the percolation order parameter. On the other hand, if $\ell \ll \xi$, the fractal irregularities of the infinite cluster will be dominant, and $\mathcal{M}(\ell) \sim \ell^{d_H}$, while being practically independent of ξ. We can therefore define $\widehat{\mathcal{M}}(x) \simeq \text{const.}$, for $x \ll 1$, which implies that $\psi = d_H$. Since $P \sim \left|p - p_c\right|^{\beta} \sim \xi^{-\beta/\nu}$, for $x \gg 1$, we have

$$\widehat{\mathcal{M}}(x) \sim x^{\beta/\nu}, \qquad \text{for } x \gg 1, \tag{11.95}$$

and therefore $\mathcal{M}(\ell) \sim \ell^{d-\beta/\nu}$. But this is compatible with expression (11.94) only if

$$d_H = d - \beta/\nu. \tag{11.96}$$

In this manner, we have related the fractal dimension of the percolation cluster to the usual critical exponents.

It would take us too long to discuss other aspects of percolation, such as its effects on conductivity, rigidity, and the like. A detailed introduction to percolation theory and its applications is due to D. Stauffer and A. Aharony [Stauf92].

11.4 Systems with Impurities

The problems pertaining to polymers and percolation are peculiar because they concern purely geometrical properties of random systems. There is, moreover, a subtle, but important, physical difference between the polymer and percolation problems. In fact, one can suppose that in the course of the observation time, a linear polymer will explore the variety of physically accessible configurations like an ordinary system in statistical mechanics.[1] In this case, ordinary statistical mechanics and thermodynamics (the expression of entropy as a function of accessible phase space volume, for instance) maintain their validity. In the case of percolation, on the contrary, the systems are composed of units of much larger than molecular scale, and once prepared, their percolative properties remain fixed. The space of percolative configurations is therefore not explored by a single system during an experiment—instead, each physical system is an *instance* of a random system extracted from a specific distribution. There is no physically interpretable entropy associated with the number of accessible configurations obtained from this distribution. Even more important is the fact that the average properties of random systems of this sort can be evaluated experimentally only by performing measures over a large number of samples, each one corresponding to a different and independent instance of the random system. In practice, for extensive properties, this average is obtained by considering as independent instances separated regions of a macroscopic sample, assuming that the random process is independently realized in each of them. Properties such as conductivity, however, are not extensive, and in this case, one can expect that there will be important fluctuations in macroscopic samples as well.

In a general disordered system, with a given preparation procedure, the details of the system's structure will differ from sample to sample. In order to produce samples such as alloys of two metals, for instance, the samples are prepared in the liquid state, at high temperature, and then rapidly cooled. In these conditions, the mutual positions of the atoms of different chemical nature do not change at room temperature during the observation time—they can be considered frozen. If the preparation temperature is high enough, it is reasonable to assume that an atom of each chemical species with a probability proportional to its concentration is present at each site of the metallic lattice, and independently for different sites. Once the sample has been prepared, this structure remains blocked.

We can describe a system of this type by means of a random Hamiltonian. Let us consider, for instance, a d-dimensional Ising system, in which a fraction p of lattice points (chosen independently from one another) contains a magnetic atom, and hosts therefore a spin variable $\sigma_i = \pm 1$, while the remaining fraction $1 - p$ of lattice points is occupied by a nonmagnetic impurity. The variables σ_i are called **fast** or **thermal** variables. The configuration $\sigma = (\sigma_i)$ of the spin variables changes in general during the observation time. On the other hand, the configuration $\tau = (\tau_i)$ of the occupation states $\tau_i \in \{0, 1\}$ of the lattice points does not change during occupation time, and is thus said to be **quenched**. Here

[1] Actually, some configurations may be kinetically inaccessible, because they might require the free ends of the polymer to knot themselves a large number of times with the rest of the polymer. In our approach, we have disregarded these details.

$\tau_i = 1$ means that site i is occupied by a magnetic atom, and $\tau_i = 0$ means that it is occupied by a nonmagnetic impurity. Thus, even if the fraction p of magnetic sites is the same, each instance of the system will have a different impurity configuration τ. One supposes that each of the τ_i, at the moment the sample is prepared, assumes a random value distributed according to a certain law. For each instance (or, as one often says, **realization**) of the disorder, we are able in principle to evaluate the thermodynamic properties by means of the usual expressions of statistical mechanics—they will, however, depend on the particular configuration of the quenched variables. The magnetization M, for instance, will be given by

$$M_\tau = \left\langle \sum_i \sigma_i \right\rangle_\tau = \frac{1}{Z_\tau} \sum_\sigma \left(\sum_i \tau_i \sigma_i \right) \mathrm{e}^{-\mathcal{H}(\sigma;\tau)/k_\mathrm{B}T}, \tag{11.97}$$

where we denote by $\tau = (\tau_i)$ the distribution of magnetic sites, the Hamiltonian $\mathcal{H}(\sigma;\tau)$ is given by

$$\mathcal{H}(\sigma;\tau) = -J \sum_{<ij>} \tau_i \sigma_i \tau_j \sigma_j - h \sum_i \tau_i \sigma_i, \tag{11.98}$$

since the interaction only acts between magnetic sites, and the partition function Z_τ is given by

$$Z_\tau = \sum_\sigma \mathrm{e}^{-\mathcal{H}(\sigma;\tau)/k_\mathrm{B}T}. \tag{11.99}$$

In a large sample, different regions will host different realizations of the disorder. Thus the observed value of an extensive observable such as M is in fact proportional to its average over the distribution of the quenched variable τ. We distinguish therefore between the **quenched** average, denoted by an overbar, performed on the distribution of the quenched variables τ, and the usual average over the fast variables, denoted by angular brackets, called the **annealed** average. We thus obtain

$$\begin{aligned} \overline{M_\tau} &= \overline{\left\langle \sum_i \sigma_i \right\rangle_\tau} = \sum_\tau P(\tau) \left\langle \sum_i \sigma_i \right\rangle_\tau \\ &= \sum_\tau P(\tau) \frac{1}{Z_\tau} \sum_\sigma \left(\sum_i \tau_i \sigma_i \right) \mathrm{e}^{-\mathcal{H}(\sigma;\tau)/k_\mathrm{B}T}, \end{aligned} \tag{11.100}$$

where $P(\tau)$ is the probability distribution of the configuration τ of the quenched variables. These expressions refer to the *quench* technique by which a metallic object is cooled rapidly in order to increase its hardness. The procedure by which a quenched metallic object is brought back to high temperature and then cooled slowly to make it more malleable is called *annealing*. This model can be straightforwardly generalized to the case in which σ_i is a vector with n components, yielding a random XY model, a Heisenberg model, and the like.

Let us try to evaluate the average magnetization per site $m = M/N$ in the mean field, by simply averaging the effective field $h_\mathrm{eff} = h + \sum_{j(\mathrm{n.n.}i)} J\tau_i m$ acting on a magnetic site i. We

obtain $h_{\text{eff}} = h + p\,\zeta\, Jm$, where ζ is the coordination number of the lattice. This yields a critical temperature $T_c(p) = p\,T_c(p{=}1)$. But this cannot be true, since when p is below the percolation threshold p_c there are no infinite clusters of interacting magnetic atoms able to sustain spontaneous magnetization. We need therefore a more powerful method.

A very popular approach is the **replica method**. It starts from the observation that, if we wish to evaluate the average of an extensive observable, we can obtain it by differentiating the average of $\ln Z_\tau$, where Z_τ is the partition function, taken over the distribution of quenched variables. Now we have the identity

$$\ln Z = \lim_{n\to 0} \frac{Z^n - 1}{n}. \tag{11.101}$$

The nth power of Z can be obtained by the expression

$$Z_\tau^n = \sum_{\sigma^1 \ldots \sigma^n} \exp\left[-\frac{1}{k_B T} \sum_{a=1}^{n} \mathcal{H}(\sigma^a; \tau) \right], \tag{11.102}$$

where the σ^a, $a = 1, \ldots, n$, are the configurations of n independent systems, *with the identical distribution τ of the impurities*. They are therefore n independent replicas of the same system, which gives the method its name.

We can now hope that it is possible to switch the summation order, averaging first over τ in order to obtain

$$\begin{aligned} \overline{Z^n} &= \sum_{\sigma^1 \ldots \sigma^n} \sum_{\tau} P(\tau) \exp\left[-\frac{1}{k_B T} \sum_{a=1}^{n} \mathcal{H}(\sigma^a; \tau) \right] \\ &= \sum_{\sigma^1 \ldots \sigma^n} \exp\left[-\frac{1}{k_B T} \mathcal{H}'(\sigma^1, \ldots, \sigma^n) \right], \end{aligned} \tag{11.103}$$

where $\mathcal{H}'(\sigma^1, \ldots, \sigma^n)$ is a nonrandom Hamiltonian that describes n magnetic lattices interacting among themselves. This expression should be then evaluated by the usual methods of statistical mechanics, such as mean-field theory or cumulant expansion, yielding a result that depends on n as a parameter. We then set

$$\overline{\ln Z} = \lim_{n\to 0} \frac{\overline{Z^n} - 1}{n}. \tag{11.104}$$

This approach has been used to set up an ϵ-expansion for the critical properties of the m-vector ferromagnet with impurities [Harr74a], whose results are described in [Lube79, sec. 2.5]. One obtains that, to first order in ϵ, the usual Heisenberg fixed point is stable when $m > 4$, and that for $m < 4$ the stable fixed point is a new one, with different values of the exponents. Thus, contrary to what one could have expected, impurities do not smear out the critical point, but rather change its behavior.

The stability of the ordinary critical point with respect to these impurities can be assessed by a heuristic argument, known as the **Harris criterion** [Harr74b]. Consider a large system as the juxtaposition of several subsystems of linear size ℓ. These systems are independent of one another if ℓ is of the order of the correlation length ξ. Then the critical temperature for each subsystem depends on the average interaction within that subsystem, and will be a random quantity. Since it arises from the average over ξ^d

random variables, its standard deviation will be of the order of $\xi^{-d/2}$. For stability, one requires that its distance from the critical temperature T_c be larger than this standard deviation. Now, since $\xi \sim |T - T_c|^{-\nu}$, this corresponds to $|T - T_c| > \xi^{-d/2} \sim |T - T_c|^{d\nu/2}$. Thus the critical point is stable against the randomness due to impurities if $d\nu > 2$, which corresponds, via the scaling law $\alpha = 2 - d\nu$, to $\alpha < 0$. This result is confirmed near $d = 4$ by the renormalization group calculations mentioned above.

A different kind of disorder is a **random field**. A random-field Ising model is described by the Hamiltonian

$$\mathcal{H}(\sigma) = -J \sum_{<ij>} \sigma_i \sigma_j - \sum_i h_i \sigma_i, \tag{11.105}$$

where the h_i are independent, identically distributed random variables for each lattice point i. It is possible to realize such a system by placing a disordered *anti*-ferromagnet in a uniform magnetic field, but the model is in fact relevant for a rich class of disordered systems, such as random alloys and binary fluid mixtures in porous media.

Let us denote the typical magnitude of the random fields h_i by h. Then, for $h \gg J$ and low temperatures, we expect the spins to align with the local field, and thus, if the distribution of h_i is even, we have $\langle \sigma_i \rangle = 0$. On the other hand, if $h \ll J$, we expect the system to behave like a perturbed ferromagnet at low temperatures. In fact, the behavior of the system is not so simple. In the small h regime, the ferromagnetic order can be unstable with respect to the overturning of large domains, as pointed out by Imry and Ma [Imry75]. Suppose that the system is in a perfectly aligned ferromagnetic state, say, pointing "up," and consider the effect of turning "down" a large droplet of linear size ℓ. The excess of down fields will be of the order of $\ell^{d/2}$, where d is the dimensionality of the system, leading to a reduction in energy of roughly $h\ell^{d/2}$. On the other hand, there is an exchange cost due to the separation of the droplet from its surroundings. For an Ising system, this is by a sharp wall, and the cost will be proportional to the surface area of the droplet, yielding a cost of roughly $J\ell^{d-1}$. Thus flipping the droplet will be favored if $d < 2$. For a Heisenberg model, or other models with the breaking of a continuous symmetry, the wall will be smeared over a length of order ℓ, yielding a cost of roughly ℓ^{d-2}. For these systems, the ferromagnetic order becomes unstable for $d < 4$.

Experimentally, systems of this kind exhibit hysteresis and an apparent inability to reach equilibrium on reasonable time scales. The nonequilibrium behavior of this and of other disordered systems is an interesting field in which research is very active, but discussing it would take us too far from our objective.

11.5 Spin Glasses

> For every complex problem there is an answer that is clear, simple, and wrong.
>
> —H. L. Mencken

A **spin glass** is a magnetic system in which the interaction among spins varies randomly from place to place and from sample to sample. Experimentally, it is prepared by

diluting the magnetic impurities in a nonmetallic magnetic matrix—for example, by diluting atoms of Mn or Fe in a matrix of Cu or Au, with a relative concentration on the order of a few percentage points. In this situation, the exchange interactions between the spins of the electrons located in the impurities and those of the conduction electrons give rise to a long-distance spin-spin interaction, whose magnitude roughly decays as r^{-3}, where r is the distance between the spin, but whose value oscillates rapidly, changing sign over distances on the order of the lattice step.

Once the sample is prepared, it is cooled to lower and lower temperatures. If the sample is cooled in the absence of a magnetic field, its magnetization vanishes, but its susceptibility to very small applied fields exhibits a well-defined cusp at a temperature of the order of a dozen Kelvin. If the magnitude of the applied field increases, the cusp rounds off rapidly.

A simple interpretation of this behavior (due to Fischer [Fisc77]) can be obtained as follows. As we know, the local susceptibility $\chi_{ij} = \partial \langle \sigma_i \rangle / \partial h_j$ of a spin placed in i with respect to a magnetic field applied in j is proportional to the spin correlation function:

$$k_B T \chi_{ij} = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma \rangle_j . \tag{11.106}$$

If one considers a given sample in a vanishing field, its global susceptibility will be given by the average of χ_{ij} over the sample. In a spin glass, we expect both the $\langle \sigma_i \rangle$ and the $\langle \sigma_i \sigma_j \rangle$ to be random quantities, roughly independent for different sites, and with a vanishing average. If this is the case, only the local terms contribute to the average susceptibility:

$$k_B T \left[\chi_{ij} \right]_{av} \simeq \left[\langle \sigma_i^2 \rangle \right]_{av} - \left[\langle \sigma_i \rangle^2 \right]_{av} , \tag{11.107}$$

where $[\ldots]_{av}$ represents the average over the sample. If we assume that $\sigma_i = \pm 1$ and that the system is in a well-defined thermodynamic state, in which $\langle \sigma_i \rangle = m_i$ (dependent on the lattice site i), this expression can be written

$$k_B T \left[\chi_{ij} \right]_{av} \simeq 1 - q_{EA}, \tag{11.108}$$

where the **Edwards-Anderson order parameter** q_{EA} is defined by

$$q_{EA} = \left[\langle \sigma_i \rangle^2 \right]_{av} . \tag{11.109}$$

Thus $[\chi]_{av} \propto T^{-1}$ as long as $q_{EA} = 0$, but behaves otherwise if $q_{EA} > 0$.

This argument may account for the experimentally observed cusp, but it does not explain a host of "strange" behaviors, especially dynamic ones, exhibited by spin glasses. More particularly, it has been realized that the hypothesis that the system is well described by an ensemble at equilibrium is not justified. Anyway, the presence of the cusp suggests that there is a temperature T_G below which the system's behavior changes radically. Understanding the nature of the thermodynamic state below this temperature is the subject of intensive research.

A standard model of a spin glass in d dimensions is the **Edwards-Anderson** model, defined on a d-dimensional lattice by the Ising-like Hamiltonian [Edwa75]

$$\mathcal{H}(\sigma) = -\sum_{<ij>} J_{ij} \sigma_i \sigma_j, \tag{11.110}$$

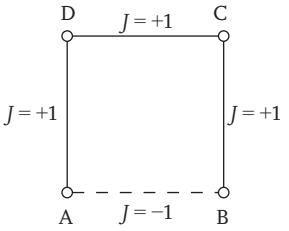


FIGURE 11.7. A frustrated plaquette in a two-dimensional Ising model with competing interactions.

where $\sigma_i = \pm 1$, $\forall i$, and the interaction constants J_{ij} relative to the nearest-neighbor pair $<ij>$ are independent, identically distributed random variables for different nearest-neighbor pairs. One often assumes J to be of the form $\pm J_0$, symmetrically distributed, or to be Gaussian-distributed with a vanishing average and a suitable variance. Although this model has been the subject of intense studies for several decades, we are still quite far from understanding its behavior. In the present section, I deal with even less realistic models, whose behavior is simpler, so as to highlight the specific phenomena that are produced by this kind of disorder.

11.5.1 *Frustration*

The presence of bonds with positive and negative interaction coefficients has an important consequence. Let us consider a two-dimensional Ising system, and focus our attention on an elementary plaquette like the one shown in figure 11.7. Obviously, a positive J_{ij} coefficient favors the parallel alignment of the spins σ_i and σ_j, while a negative coefficient favors their antiparallel alignment. Now, it is easy to see that in the situation illustrated, it is not possible to find an arrangement of the four spins placed at the plaquette's vertices that satisfies all four bonds. In effect, if we, for example, decide to arrange all the spins up, the AB bond, where J is negative, will not be satisfied; if we instead choose, for example, to arrange the spin located in A or B down, so as to satisfy that bond, then either the AD or the BC bond, respectively, will not be satisfied. This phenomenon is called **frustration**, and the plaquette is called **frustrated**. This situation takes place when the product of the J_{ij}s over all the plaquette's bonds is negative:

$$J_{AB}J_{BC}J_{CD}J_{DA} < 0. \tag{11.111}$$

In the presence of frustration, the minimal energy states tend to be very numerous, because it is often possible to pass from one state to another with a similar value of the energy by simply changing the choice of which bonds not to satisfy. Consider a two-dimensional Ising model with interactions of both signs. The energy cost of a configuration is proportional to the total number of unsatisfied bonds, that is, positive bonds connecting antiparallel spins or negative bonds connecting parallel ones. These bonds form barriers in the lattice. In a system without frustration, as we have seen in the Peierls argument (sec. 5.9), the barriers close on themselves. When frustration is present some barriers connect pairs of frustrated plaquettes. Thus the problem of finding the minimum energy configuration boils down to that of connecting frustrated plaquettes on a lattice

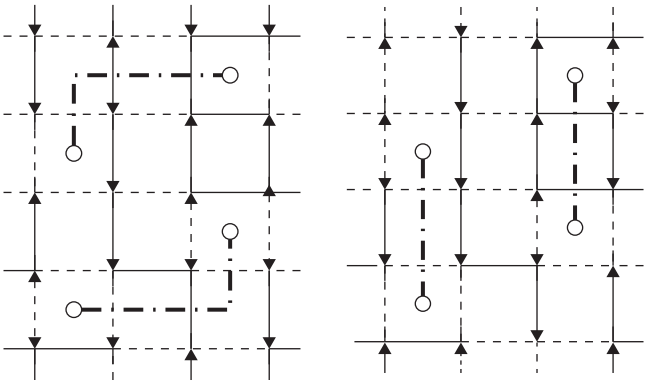


FIGURE 11.8. Two examples of minimal-length barrier configurations in a square lattice with frustration. Full lines represent bonds with $J > 0$, broken lines those with $J < 0$. Four frustrated plaquettes are shown, marked by the circles. The dash-dotted lines represent a possible minimal-energy configuration of the frustrated bonds. The arrows represent a possible spin configuration corresponding to the given pattern of frustrated bonds. The two configurations differ by the position of the frustrated barriers.

with barriers (paths in the dual lattice) of minimal length (see fig. 11.8). This is a nontrivial combinatorial problem. Frustration is therefore often associated with situations where there is a large number of microscopic states with minimal or close to minimal energy, quite different from one another.

Frustration does not occur only in random systems. For example, in an antiferromagnetic two-dimensional Ising model, in which the spins are arranged over a triangular lattice, all the cells are frustrated. In this case, it has been proved that the system does not exhibit phase transitions at a finite temperature with a vanishing magnetic field, while it exhibits particular forms of order if the magnetic field breaks the degeneration of the ground state caused by the frustration.

On the other hand, the simple appearance of interactions of different signs does not necessarily lead to frustration. One example is the **Mattis model** [Matt76], that is, an infinite-range Ising model in which the interaction J_{ij} between spin i and spin j is given by

$$J_{ij} = J_0 \xi_i \xi_j, \tag{11.112}$$

where $\xi = (\xi_i)$ is a collection of N independent random variables, equal to $+1$ or -1 with equal probability. Clearly, the minimal energy state corresponds to $\sigma_i = \xi_i$, $\forall i$, or to the opposite configuration.

11.5.2 The Random-Energy Model

We can attempt to achieve a better understanding of the effects of the interaction disorder by considering a simple solvable model in which the energy levels are independent random variables. Let us consider a system composed of N Ising spins $\sigma_i = \pm 1$, having thus 2^N different configurations $\sigma = (\sigma_i)$, $i = 1, \ldots, N$. In the random-energy model (REM), introduced by B. Derrida [Derr81], the energy $E(\sigma)$ of each configuration σ is an independent,

identically distributed random variable with distribution $P(E)$. In what follows, we assume that $P(E)$ is a Gaussian distribution with vanishing mean and variance equal to $NJ_0^2/2$:

$$P(E) \propto \exp\left(-\frac{E^2}{NJ_0^2}\right). \tag{11.113}$$

In this way, the typical values of the energy are of order NJ_0. Let us now consider the number $\omega(E)\,\mathrm{d}E$ of states whose energy lies between E and $E+\mathrm{d}E$. This number fluctuates from one sample to the next, but its average $\overline{\omega(E)}\,\mathrm{d}E$ over the distribution $P(E)$ can be easily calculated:

$$\overline{\omega(E)}\,\mathrm{d}E = 2^N\,P(E)\,\mathrm{d}E = 2^N\left(N\pi J_0^2\right)^{-1/2}\exp\left(-\frac{E^2}{NJ_0^2}\right)\mathrm{d}E. \tag{11.114}$$

Let us choose $\mathrm{d}E$ to be proportional to $N^{1/2}$. We then have

$$\overline{\omega(E)}\,\mathrm{d}E \propto \exp\left(N\ln 2 - \frac{E^2}{N\,J_0^2}\right). \tag{11.115}$$

Then for $|E| < N\,J_0\sqrt{\ln 2} = E_0$ the number of states in the given energy interval grows exponentially with N. In this regime, the entropy is given by

$$S(E) = k_\mathrm{B}\ln\overline{\omega(E)}\,\mathrm{d}E \simeq Nk_\mathrm{B}\left(\ln 2 - \frac{E^2}{N^2\,J_0^2}\right). \tag{11.116}$$

The entropy has a maximum for $E=0$, which corresponds to $T=\infty$. Thus for positive values of E the corresponding temperature is negative (see sec. 3.11). We are thus interested in *negative* values of E. The temperature T is given by

$$\frac{1}{T} = \frac{\mathrm{d}S}{\mathrm{d}E} = -k_\mathrm{B}\frac{2E}{N\,J_0^2}. \tag{11.117}$$

Using these expressions, we obtain $E(T) = -NJ_0^2/2k_\mathrm{B}T$, and the free energy per degree of freedom $f = (E-TS)/N = -k_\mathrm{B}T(\ln 2 - J_0^2/k_\mathrm{B}T)$. This is all right as long as the density of states $\overline{\omega(E)}$ is exponentially large in N. But when $E < -E_0$ this expression becomes negative: this is not acceptable, since the number of states available to the system cannot be less than 1. Indeed, this result means that the density of states *decreases* exponentially with N, and therefore that in most samples there are no states with the given value of the energy. Therefore, at temperatures smaller than T_0, where

$$\frac{1}{T_0} = \left.\frac{\mathrm{d}S}{\mathrm{d}E}\right|_{-E_0} = k_\mathrm{B}\frac{2\ln 2}{J_0}, \tag{11.118}$$

the system can only remain frozen in one of the few available low-energy states, with energy close to $-E_0$, and its entropy vanishes.

However, this does not mean that for $T < T_0$, the system will always be found in a single well-defined state, say, of minimal energy. We are dealing in fact with a phase coexistence, in which each of the coexisting phases has vanishing entropy, because the system occupies

a number of states that grows with N more slowly than exponentially. We discuss this coexistence using an approach developed by Bouchaud and Mézard [Bouc97].

The partition function of the model is given by

$$Z = \sum_{\alpha=1}^{M} z_\alpha, \tag{11.119}$$

where α denotes the state, $M = 2^N$ is the total number of states, and the Boltzmann factor z_α is given by $z_\alpha = \exp(-E_\alpha/k_B T)$, where the energies E_α are distributed according to $P(E)$. We are interested in the tail of the distribution, when the E_α are large and negative. Let us discuss the distribution of the minimum energy E^*. Introducing the cumulative distribution function $\Phi(E)$, defined by

$$\Phi(E) = \int_{-\infty}^{E} \mathrm{d}E'\, P(E'), \tag{11.120}$$

the distribution $P_M(E^*)$ has the expression

$$P_M(E^*) = MP(E^*)\,[1-\Phi(E^*)]^{M-1} = -\left.\frac{\mathrm{d}}{\mathrm{d}E}\,[1-\Phi(E)]^M\right|_{E^*}. \tag{11.121}$$

For large values of M, we have

$$[1-\Phi(E)]^M \simeq \exp\left[-M\Phi(E)\right]. \tag{11.122}$$

In correspondence with the critical value of the energy $E_c = -E_0 = -NJ_0\sqrt{\ln 2}$, the cumulative distribution function of E becomes very small, $M\Phi(E_c) \simeq 1$, meaning that there are a finite number of states below E_c, even if N is large. Let us set $E^* = E_c + \epsilon$, with $|\epsilon| \ll |E_c|$. We then have, to first order in ϵ, and taking into account that $M\Phi(E_c) \simeq 1$,

$$\begin{aligned}
\exp\left[-M\Phi(E)\right] &= \exp\left\{-M\exp\left[\ln\Phi(E)\right]\right\} \\
&\simeq \exp\left\{-M\exp\left[\ln\Phi(E_c) + \left.\frac{\mathrm{d}\ln\Phi}{\mathrm{d}E}\right|_{E_c}\epsilon\right]\right\} \\
&= \exp\left[-\exp\left(\frac{2E_c}{NJ_0^2}\epsilon\right)\right] = \exp\left[-\exp\left(\frac{2\sqrt{\ln 2}}{J_0}\,\epsilon\right)\right].
\end{aligned} \tag{11.123}$$

Defining the rescaled variable $u = 2\sqrt{\ln 2}\,\epsilon/J_0$, and taking the derivative, we obtain the distribution of u:

$$P^*(u) = \exp\left(u - \exp u\right). \tag{11.124}$$

This distribution of the extreme values of a random variable possesses a much more general validity, and is called the **Gumbel distribution**. It is shown in fig. 11.9. It decreases exponentially for negative values of u, and much faster for positive ones. It exhibits a maximum for $u = 0$, which confirms that E_c is the most probable minumum energy value.

Exercise 11.17 Show that the Gumbel distribution holds for the minimum value of a collection of M random variables E_α, identically distributed, whose distribution for large and negative E is given by $P(E) \sim \exp\left[-B\,|E|^\delta\right]$, with $B, \delta > 0$.

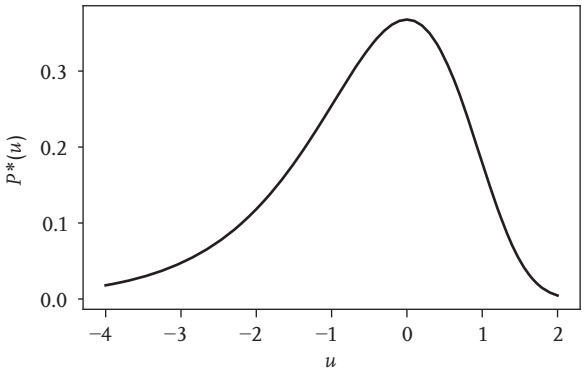


FIGURE 11.9. The Gumbel distribution, governing the rescaled deviation $u = 2\sqrt{\ln 2}\,\epsilon/J_0$ of the minimal energy value from E_c.

If the E_α are distributed approximately exponentially for large and negative E, the Boltzmann factors z_α decay as a power law:

$$P_z(z) \sim z^{-1-\mu}, \tag{11.125}$$

where

$$\mu = \frac{T}{T_0}, \tag{11.126}$$

where T_0 is given by eq. (11.118).

Exercise 11.18 Derive this relation.

For $\mu > 1$, all terms in the partition function contribute a small amount. For $\mu < 1$, that is, $T < T_0$, the average of Z diverges, and only a small number of terms contributes. In order to characterize this behavior, we introduce the relative weights,

$$w_\alpha = \frac{z_\alpha}{Z}. \tag{11.127}$$

Assuming $T < T_0$, we have

$$w_\alpha = \frac{z_\alpha}{z_\alpha + Z'_\alpha}, \tag{11.128}$$

where $Z'_\alpha = \sum_{\beta\,(\neq\alpha)} z_\beta$ is independent of z_α and of order $M^{1/\mu}$. Given a value w of the weight, denote by $\zeta(w)$ the corresponding value of z. Then the distribution of the weights with a fixed value of Z'_α is given by

$$\begin{aligned} P_w(w) &= \int \mathrm{d}z\, \delta\left(w - \frac{z}{z+Z'}\right) P_z(z) = \left.\frac{\mathrm{d}z}{\mathrm{d}w}\right|_{w(z)} P_z(\zeta(w)) \\ &= \frac{Z'_\alpha}{(1-w)^2} P_z\left(\frac{Z'w}{1-w}\right). \end{aligned} \tag{11.129}$$

In order to have a nonvanishing weight w in the limit of large M, z must be large. We can therefore apply the asymptotic form for the distribution of z, thus obtaining, for $w \gg M^{-1/\mu}$,

$$P_w(w) = \frac{C}{M}(1-w)^{\mu-1} w^{-\mu-1}, \tag{11.130}$$

where the constant C is fixed by the condition $\overline{w} = \int_0^1 \mathrm{d}w\, w P_w(w) = 1/M$. Using eq. (E.15), we obtain $1/C = \Gamma_{\mathrm{E}}(\mu)\Gamma_{\mathrm{E}}(1-\mu)$, where $\Gamma_{\mathrm{E}}(z)$ is Euler's gamma function. We can thus calculate the moments $\overline{w^k}$, with $k > 1$, always via eq. (E.15), and we obtain, for $\mu < 1$,

$$\overline{w^k} = M \int_0^1 \mathrm{d}w\, w^k P_w(w) = \frac{\Gamma_{\mathrm{E}}(k-\mu)}{\Gamma_{\mathrm{E}}(k)\Gamma_{\mathrm{E}}(1-\mu)}. \tag{11.131}$$

In particular, one obtains

$$\overline{w^2} = 1 - \mu. \tag{11.132}$$

Now, $1/\overline{w^2}$ measures the typical number of states that contribute to the partition function. We see that this quantity is finite for $\mu < 1$ and diverges as $|T - T_0|^{-1}$ as T approaches T_0 from below. The average energy per degree of freedom of the system is constant in the low-temperature phase and is given by

$$\frac{\overline{E}}{N} = \frac{E_{\mathrm{c}}}{N} + \langle u \rangle \frac{J_0}{2N}\sqrt{\ln 2}. \tag{11.133}$$

Thus the energy per degree of freedom has fluctuations of order $1/N$, distributed according to the Gumbel distribution. The average $\langle u \rangle$ is equal to $\Gamma'_{\mathrm{E}}(1) = -\gamma_{\mathrm{E}} \simeq -0.577216$, where γ_{E} is Euler's gamma constant. We conclude that, at temperatures below T_0, the system remains frozen in one of its low-energy states, whose energy exhibits fluctuations of order 1, distributed according to the Gumbel distribution.

11.5.3 The Random-Energy Model as Limit of a Spin Model

The definition of the random-energy model is very abstract and can leave one puzzled. One can, however, obtain it as the limit of a mean-field spin model of a more familiar kind.

Given a system of N spins $\sigma = (\sigma_i)$, $i = 1, \ldots, N$, define the p-spin Hamiltonian by

$$\mathcal{H}_J^{(p)}(\sigma) = -\sum_{\{i_1,\ldots,i_p\}} J_{i_1,\ldots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \tag{11.134}$$

where the sum runs over all the sets $\{i_1, \ldots, i_p\}$ of p distinct indices. The coefficients $J_{i_1,\ldots,i_p}$ are random variables, independent for different sets $\{i_1, \ldots, i_p\}$, identically distributed with the Gaussian distribution

$$P_J(J) = \sqrt{\frac{N^{p-1}}{\pi\, p!}} \exp\left(-\frac{J^2\, N^{p-1}}{J_0^2\, p!}\right). \tag{11.135}$$

The power of N is chosen so as to guarantee a good thermodynamic limit.

We now want to show that, in a suitably defined limit $p \to \infty$, this Hamiltonian describes a random-energy model like the one we just discussed. First, let us evaluate the probability distribution $P(E)$ of the energy of a given spin configuration σ:

$$P(E) = \int \prod_{\{i_1,\ldots,i_p\}} \mathrm{d}J_{\{i_1,\ldots,i_p\}}\, P_J\left(J_{\{i_1,\ldots,i_p\}}\right)\, \delta\left(E - \mathcal{H}_J^{(p)}(\sigma)\right). \tag{11.136}$$

We can make use of the Fourier representation of the delta function and evaluate the resulting integral by the saddle point method. We thus have

$$\begin{aligned} P(E) &= \int \frac{\mathrm{d}\lambda}{2\pi\mathrm{i}} \int \prod_{\{i_1,\ldots,i_p\}} \mathrm{d}J_{\{i_1,\ldots,i_p\}}\, P_J\left(J_{\{i_1,\ldots,i_p\}}\right) \\ &\qquad \times \exp\left(\mathrm{i}\lambda E + \mathrm{i}\lambda \sum_{\{i_1,\ldots,i_p\}} J_{i_1,\ldots,i_p}\sigma_{i_1}\cdots\sigma_{i_p}\right) \\ &= \int \frac{\mathrm{d}\lambda}{2\pi\mathrm{i}}\, \exp\left(\mathrm{i}\lambda E - \frac{\lambda^2 J_0^2 p!}{4N^{p-1}} \sum_{\{i_1,\ldots,i_p\}} \sigma_{i_1}^2\cdots\sigma_{i_p}^2\right), \end{aligned} \tag{11.137}$$

where we have evaluated the Gaussian integral over the Js. Obviously, $\sigma_i^2 = 1, \forall i$, and therefore, the terms under summation in this expression are all equal. It is easy to see that there are $N(N-1)\cdots(N-p+1)/p! \simeq N^p/p!$ such terms. We thus obtain

$$P(E) = \int \frac{\mathrm{d}\lambda}{2\pi\mathrm{i}}\, \exp\left(\mathrm{i}\lambda E - \frac{\lambda^2 N J_0^2}{4}\right), \tag{11.138}$$

and calculating the Gaussian integral one more time, we finally obtain

$$P(E) \propto \exp\left(-\frac{E^2}{NJ_0^2}\right). \tag{11.139}$$

Let us now evaluate the joint distribution $P(E_1, E_2)$ of the energies of two different configurations $\sigma^{(1)}$ and $\sigma^{(2)}$:

$$P(E_1, E_2) = \overline{\delta\left(E_1 - \mathcal{H}_J^{(p)}(\sigma^{(1)})\right)\delta\left(E_2 - \mathcal{H}_J^{(p)}(\sigma^{(2)})\right)}. \tag{11.140}$$

We obtain

$$\begin{aligned} P(E_1, E_2) &= \int \frac{\mathrm{d}\lambda_1}{2\pi\mathrm{i}} \int \frac{\mathrm{d}\lambda_2}{2\pi\mathrm{i}} \int \prod_{\{i_1,\ldots,i_p\}} \mathrm{d}J_{\{i_1,\ldots,i_p\}}\, P\left(J_{\{i_1,\ldots,i_p\}}\right) \\ &\quad \times \exp\left(\mathrm{i}\lambda_1 E_1 + \mathrm{i}\lambda_2 E_2\right) \\ &\quad \times \exp\left[\sum_{\{i_1,\ldots,i_p\}} J_{i_1,\ldots,i_p}\left(\mathrm{i}\lambda_1 \sigma_{i_1}^{(1)}\cdots\sigma_{i_p}^{(1)} + \mathrm{i}\lambda_2 \sigma_{i_1}^{(2)}\cdots\sigma_{i_p}^{(2)}\right)\right]. \end{aligned} \tag{11.141}$$

By evaluating the Gaussian integral over the Js, we obtain

$$P(E_1, E_2) = \int \frac{\mathrm{d}\lambda_1}{2\pi\,\mathrm{i}} \int \frac{\mathrm{d}\lambda_2}{2\pi\,\mathrm{i}}\, \exp\left(\mathrm{i}\lambda_1 E_1 + \mathrm{i}\lambda_2 E_2\right) \times \exp\left[-\frac{J_0^2 p!}{4N^{p-1}} \sum_{\{i_1,\ldots,i_p\}} \left(\lambda_1 \sigma_{i_1}^{(1)} \cdots \sigma_{i_p}^{(1)} + \lambda_2 \sigma_{i_1}^{(2)} \cdots \sigma_{i_p}^{(2)}\right)^2 \right]. \tag{11.142}$$

The quantity under summation can be written

$$\lambda_1^2 + \lambda_2^2 + 2\lambda_1\lambda_2 \left(\sigma_{i_1}^{(1)} \sigma_{i_1}^{(2)}\right) \cdots \left(\sigma_{i_p}^{(1)} \sigma_{i_p}^{(2)}\right).$$

Let us introduce the **overlap** $q\left(\sigma^{(1)}, \sigma^{(2)}\right)$ between the configurations $\sigma^{(1)}$ and $\sigma^{(2)}$:

$$q\left(\sigma^{(1)}, \sigma^{(2)}\right) = \frac{1}{N} \sum_i \sigma_i^{(1)} \sigma_i^{(2)}. \tag{11.143}$$

This quantity is equal to 1 if $\sigma^{(1)}$ and $\sigma^{(2)}$ differ by a finite number of spins and is otherwise smaller than 1. If $\sigma^{(1)}$ and $\sigma^{(2)}$ are opposites, $q\left(\sigma^{(1)}, \sigma^{(2)}\right) = -1$. Therefore, q is a good measure of the resemblance between spin configurations and plays a fairly important role in the theory.

A little algebra shows that, up to terms negligible for large N, one has

$$\sum_{\{i_1,\ldots,i_p\}} \left(\sigma_{i_1}^{(1)} \sigma_{i_1}^{(2)}\right) \cdots \left(\sigma_{i_p}^{(1)} \sigma_{i_p}^{(2)}\right) = \frac{N^p}{p!} \left[q\left(\sigma^{(1)}, \sigma^{(2)}\right)\right]^p. \tag{11.144}$$

We thus obtain

$$P(E_1, E_2) = \int \frac{\mathrm{d}\lambda_1}{2\pi\,\mathrm{i}} \int \frac{\mathrm{d}\lambda_2}{2\pi\,\mathrm{i}}\, \exp\left(\mathrm{i}\lambda_1 E_1 + \mathrm{i}\lambda_2 E_2\right) \times \exp\left[-\frac{N J_0^2}{4} \left(\lambda_1^2 + \lambda_2^2 + 2\lambda_1\lambda_2\, q^p\right) \right]. \tag{11.145}$$

Let us consider two configurations that are neither identical nor opposite, such that $|q| < 1$. If we take the $p \to \infty$ limit of this expression, the term proportional to $\lambda_1\lambda_2$ vanishes, and therefore $P(E_1, E_2)$ factorizes into $P(E_1) \times P(E_2)$—in other words, in this limit, the energies of different configurations are independent random variables. Thus the random-energy model is the limit, for $p \to \infty$, of a p-spin model like the one considered in this subsection.

11.5.4 Replica Method Approach to the Random-Energy Model

The same results can be obtained by the replica method, introduced in sec. 11.4. The method requires us to evaluate the moments $\overline{Z^n}$ of the partition function for arbitrary real values of n, in order to obtain $\overline{\ln Z}$ by taking the $n \to 0$ limit of $(\overline{Z^n} - 1)/n$. In our case, we start from the cumulant generating function of the energy distribution, defined by

$$g(\lambda) = \frac{1}{N} \ln \int_{-\infty}^{+\infty} \mathrm{d}E\, P(E)\, \mathrm{e}^{-\lambda E}, \tag{11.146}$$

where the factor $1/N$ is introduced for convenience. A straightforward Gaussian integration yields

$$g(\lambda) = \left(\frac{\lambda J_0}{2}\right)^2 . \tag{11.147}$$

The nth moment of Z is given by

$$\overline{Z^n} = \overline{\sum_{i_1,\ldots,i_n} z_{i_1} \cdots z_{i_n}} = \overline{\sum_{i_1,\ldots,i_n} \exp\left(-\frac{1}{k_B T} \sum_i E_i \sum_{a=1}^{n} \delta_{ii_a}\right)}. \tag{11.148}$$

Using the definition (11.146), we can easily evaluate the average over $P(E)$, obtaining

$$\overline{Z^n} = \sum_{i_1,\ldots,i_n} \exp\left[N \sum_i g\left(\frac{1}{k_B T} \sum_{a=1}^{n} \delta_{ii_a}\right)\right]. \tag{11.149}$$

Let us suppose that the sum is dominated by configurations in which all the indices $i_1, \ldots, i_n$ are different. We then obtain

$$\begin{aligned} \overline{Z^n} &= M(M-1)\cdots(M-n+1) \exp\left[n\, N\, g\left(\frac{1}{k_B T}\right)\right] \\ &\simeq \exp\left\{n\left[\ln M + N\, g\left(\frac{1}{k_B T}\right)\right]\right\}. \end{aligned} \tag{11.150}$$

Taking the $n \to 0$ limit, we obtain the free energy per degree of freedom:

$$f = -\frac{k_B T}{N} \overline{\ln Z} = -k_B T\, \ln 2 - \frac{J_0^2}{4 k_B T} = f_0(T), \tag{11.151}$$

which corresponds to the expression obtained in sec. 11.5.2.

However, this result cannot hold at all temperatures, since the predicted entropy per particle, given by $s_0(T) = -\mathrm{d}f_0/\mathrm{d}T$, has the expression

$$s_0(T) = k_B \left(\ln 2 - \frac{J_0^2}{4(k_B T)^2}\right), \tag{11.152}$$

and becomes negative for $T < T_0 = J_0/(2k_B\sqrt{\ln 2})$, that is, below the transition temperature that we had evaluated in sec. 11.5.2.

For $T < T_0$ we cannot treat all the replicas on an equal footing. This means that the permutation symmetry of all replicas is spontaneously broken. This is known as **replica symmetry breaking** (RSB). The simplest RSB scheme is one in which the n replicas are divided into n/m goups of m equal indices, according to the scheme

$$\begin{aligned} i_1 = i_2 = \cdots = i_m &= k_1, \\ i_{m+1} = i_{m+2} = \cdots = i_{2m} &= k_2, \\ &\vdots \\ i_{n-m+1} = i_{n-m+1} = \cdots = i_n &= k_{n/m}, \end{aligned}$$

in which the indices $k_1, \ldots, k_{n/m}$ are different from one another. We then obtain

$$\begin{aligned}\overline{Z^n} &= M(M-1)\cdots(M-n/m+1)\left(\frac{n!}{(m!)^{n/m}}\right)\exp\left[\frac{n}{m}Ng\left(\frac{m}{k_{\rm B}T}\right)\right] \\ &\simeq \exp\left\{\frac{n}{m}\left[\ln M - m\ln\frac{m}{n} + Ng\left(\frac{m}{k_{\rm B}T}\right)\right]\right\}.\end{aligned} \tag{11.153}$$

Taking the $n\to 0$ limit and dividing by N, we obtain

$$f(T) = f_0\left(\frac{T}{m}\right), \tag{11.154}$$

where $f_0(T)$ is defined in eq. (11.151).

The extremum of this free energy is reached for

$$0 = \frac{\partial f}{\partial m} = \frac{\partial}{\partial m} f_0\left(\frac{T}{m}\right) = \frac{k_{\rm B}T\ln 2}{m^2} - \frac{J_0^2}{k_{\rm B}T}, \tag{11.155}$$

which yields

$$m = \frac{T}{T_0} = \mu, \tag{11.156}$$

where T_0 is the transition temperature defined in eq. (11.118) and μ is defined in eq. (11.126). We may check that this corresponds to a local *maximum* of f, rather than to a minimum or an inflection point. This is just one of the many surprises that appear in the replica method. As we see, the entropy vanishes for this solution, and therefore the system remains frozen for $T < T_0$.

We can also evaluate the moments $\overline{w^k}$, obtaining

$$\begin{aligned}\overline{w^k} &= \sum_i \overline{\left(\frac{z_i}{Z}\right)^k} = \lim_{n\to 0}\overline{\sum_i z_i^k Z^{n-k}} \\ &= \lim_{n\to 0}\frac{1}{n(n-1)\cdots(n-k+1)}\sum_{a_1,\ldots,a_k}\sum_{i_1,\ldots,i_n}\overline{z_{i_1}\cdots z_{i_n}}\prod_{j=1}^{k}\delta_{i_{a_1}i_{a_j}},\end{aligned} \tag{11.157}$$

where the sum over the as runs over different replica indices (between 1 and n), but the Kronecker delta product forces these replicas to have identical configurations. This implies that the k replica indices $a_1, \ldots, a_k$ must belong to the same group of m indices. Once a_1 has been chosen, this can be performed in $(m-1)\cdots(m-k+1)$ different ways. Therefore,

$$\overline{w^k} = \lim_{n\to 0}\frac{m(m-1)\cdots(m-k+1)}{n(n-1)\cdots(n-k+1)}\,\overline{Z^n}. \tag{11.158}$$

Note that

$$\lim_{n\to 0}(n-1)\cdots(n-k+1) = (-1)^{k-1}(k-1)! = (-1)^{k-1}\Gamma_{\rm E}(k), \tag{11.159}$$

while

$$(m-1)\cdots(m-k+1) = (-1)^{k-1}(1-m)\cdots(k-1-m) = (-1)^{k-1}\frac{\Gamma_{\rm E}(k-m)}{\Gamma_{\rm E}(1-m)}, \tag{11.160}$$

where we have exploited the relation $z\Gamma_{\rm E}(z) = \Gamma_{\rm E}(z+1)$. Using these relations and the fact that $m = \mu$, we obtain

$$\overline{w^k} = \frac{\Gamma_{\rm E}(k-\mu)}{\Gamma_{\rm E}(k)\Gamma_{\rm E}(1-\mu)}, \tag{11.161}$$

which agrees with the direct calculation. Thus the behavior of the random-energy model is described by a replica theory with the replica symmetry breaking (RSB). What is its interpretation? Since $\overline{w^2} = 1-\mu$, if we pick up two identical replicas of the system at the temperature $T < T_0$, they will be found in the same configuration with probability $1-\mu = 1 - T/T_0$. This corresponds to phase coexistence, in which each phase contains a single configuration.

Exercise 11.19 (Large deviations) Let $\mathcal{N}(\epsilon, \delta)$ be the total number of configurations j such that $N\epsilon < \left|E_j\right| < N(\epsilon+\delta)$, with $|\epsilon| > \epsilon_* = \sqrt{\ln 2}$ and $\delta > 0$. Show that the fraction of samples in which there exist such configurations is exponentially small.

Exercise 11.20 From the original motivation of the REM as a simple version of a spin glass, one can define a generalization of the REM in the presence of a magnetic field B. The 2^N configurations are divided into $N+1$ groups. Each group is labeled by its "magnetization" $M \in \{-N, -N+2, \ldots, N-2, N\}$, and includes $\binom{N}{(N+M)/2}$ configurations. Their energies $\{E_j\}$ are independent Gaussian variables with variance $\sqrt{N/2}$ and mean $\overline{E_j} = -MB$, which depend upon the group j belongs to. Show that there exists a phase transition line $\beta_{\rm c}$ in the plane(β, B) such that:

$$\frac{1}{N}\overline{\langle M\rangle} = \begin{cases} \tanh(\beta B), & \text{for } \beta \le \beta_{\rm c}(B); \\ \tanh(\beta_{\rm c}(B) B), & \text{for } \beta > \beta_{\rm c}(B). \end{cases}$$

Plot the magnetic susceptibility

$$\chi = \left.\overline{\frac{\partial \langle M\rangle}{\partial B}}\right|_{B=0}$$

vs. $T = 1/\beta$.

Exercise 11.21 Consider a generalization of the REM where the distribution of energies, instead of being Gaussian, is $P(E) \propto {\rm e}^{-C|E|^\delta}$, where $\delta > 0$. Show that, in order to have extensive thermodynamic potentials, one should scale C as $C = N^{1-\delta}\hat{C}$ (i.e., the thermodynamic limit $N \to \infty$ should be taken at fixed $\hat{C}$). Compute the critical temperature and the ground-state energy density. What is the qualitative difference between the cases $\delta > 1$ and $\delta < 1$?

11.5.5 The Sherrington-Kirkpatrick Model

The Sherrington-Kirkpatrick model [Sher75] is the paradigmatic model of a spin glass. It is the mean-field version of the Edwards-Anderson model. It is described by the

Hamiltonian

$$\mathcal{H}(\sigma) = -\sum_{(ij)} J_{ij}\sigma_i\sigma_j - h\sum_i \sigma_i, \tag{11.162}$$

where $\sigma = (\sigma_i)$, $\sigma_i = \pm 1$, $i = 1, \ldots, N$, and, for each pair (ij) of indices, $J_{ij} = J_{ji}$ is an independent random variable, distributed according to

$$P(J) = \sqrt{\frac{N}{2\pi J_0^2}} \exp\left(-\frac{NJ^2}{2J_0^2}\right). \tag{11.163}$$

This model was studied by several researchers over a few years, and finally solved by G. Parisi, who introduced a sophisticated RSB scheme [Pari79]. This solution earned him the 2021 Nobel Prize in physics. In this section, I sketch the difficulties of a simple-minded replica approach to this model and describe the grand structure of Parisi's solution.

To simplify the notation, we measure temperature in units of $T_0 = J_0/k_B$ and magnetic field in units of J_0. Then, by a straightforward Gaussian integration over the Js, we obtain the following expression of the nth moment of the partition function Z:

$$\begin{aligned}\overline{Z_J^n} &= \int \prod_{(ij)} \left[\mathrm{d}J_{ij}\, P(J_{ij})\right] \sum_{\sigma^1,\ldots,\sigma^n} \exp\left[\frac{1}{T}\sum_{a=1}^{n}\left(\sum_{(ij)} J_{ij}\sigma_i^a\sigma_j^a + h\sum_i \sigma_i^a\right)\right] \\ &= \sum_{\sigma^1,\ldots,\sigma^n} \exp\left[\frac{1}{2NT^2}\sum_{(ij)}\left(\sum_{a=1}^{n}\sigma_i^a\sigma_j^a\right)^2 + \frac{h}{T}\sum_a\sum_i \sigma_i^a\right] \\ &= \sum_{\sigma^1,\ldots,\sigma^n} \exp\left[\frac{Nn}{4T^2} + \frac{N}{2T^2}\sum_{(ab)}\left(\frac{1}{N}\sum_i \sigma_i^a\sigma_i^b\right)^2 + \frac{h}{T}\sum_a\sum_i \sigma_i^a\right].\end{aligned} \tag{11.164}$$

Here $\sum_{(ab)}$ denotes the sum over all distinct pairs of replicas, and we identify in the summand the square of the overlap q^{ab} between the two replicas a and b.

We now exploit the identity

$$\sqrt{\frac{\alpha^2}{2\pi}} \int \mathrm{d}x\, \mathrm{e}^{-\alpha^2x^2/2 + \alpha^2 x\gamma} = \mathrm{e}^{\alpha^2\gamma^2/2} \tag{11.165}$$

to get rid of the term proportional to $(\sum_i \sigma_i^a\sigma_i^b)^2$. Introducing the dummy integration variables Q_{ab}, for each distinct pair (a, b), $a, b \in \{1, \ldots, n\}$, we obtain

$$\begin{aligned}\exp\left[\frac{N}{2T^2}\sum_{(ab)}\left(\frac{1}{N}\sum_i \sigma_i^a\sigma_i^b\right)^2\right] &= \int \prod_{(ab)}\left(\mathrm{d}Q_{ab}\sqrt{\frac{N}{2\pi T^2}}\right) \\ &\times \exp\left[-\frac{N}{2T^2}\sum_{(ab)} Q_{ab}^2 + \frac{1}{2T^2}\sum_{(ab)} Q_{ab}\sum_i \sigma_i^a\sigma_i^b\right].\end{aligned} \tag{11.166}$$

One now has

$$\sum_{\sigma^1,\ldots,\sigma^n} \exp\left(\frac{1}{2T^2}\sum_{(ab)} Q_{ab}\sum_i \sigma_i^a\sigma_i^b + \frac{h}{T}\sum_a\sum_i \sigma_i^a\right) = \left[\sum_{s^1,\ldots,s^n} \exp\left(\frac{1}{2T^2}\sum_{(ab)} Q_{ab}\, s^a s^b + \frac{h}{T}\sum_a s^a\right)\right]^N, \tag{11.167}$$

where the variables s^a, $a = 1,\ldots,n$, take the values ± 1, since the exponential factorizes into independent sums for each index i. We thus obtain

$$\overline{Z_J^n} = \int \prod_{(ab)}\left(\mathrm{d}Q_{ab}\sqrt{\frac{N}{2\pi T^2}}\right) \times \exp\left(\frac{Nn}{4T^2} - \frac{N}{2T^2}\sum_{(ab)} Q_{ab}^2 + N\ln Z(Q)\right), \tag{11.168}$$

where

$$Z(Q) = \sum_{s^1,\ldots,s^n} \exp\left[\frac{1}{T}\left(\frac{1}{T}\sum_{(ab)} Q_{ab}\, s^a s^b + h\sum_a s^a\right)\right] = \sum_{s^1,\ldots,s^n} \exp\left[-\frac{\mathcal{H}_Q(s)}{T}\right]. \tag{11.169}$$

Thus $Z(Q)$ is the partition function of the single-site Hamiltonian $\mathcal{H}_Q(s)$, defined on the collection $s = (s^a)$ of n spins, and with the interaction matrix (Q_{ab}/T).

The expression (11.168) of $\overline{Z_J^n}$ that we have obtained suggests the use of saddle-point integration. We have therefore to identify the extrema of the expression in brackets, which we denote by $-N\,A(Q)$, where

$$A(Q) = -\frac{n}{4T^2} + \frac{1}{2T^2}\sum_{(ab)} Q_{ab}^2 - \ln Z(Q). \tag{11.170}$$

One can notice that the extremum condition $\partial A/\partial Q_{ab} = 0$ corresponds to

$$Q_{ab} = \left\langle s^a s^b\right\rangle, \tag{11.171}$$

where the average is defined with respect to the single-site Hamiltonian $\mathcal{H}_Q(s)$. Thus Q_{ab} can be interpreted as the average overlap between replica a and replica b.

To find the extrema, we have to parametrize in some form the matrix Q. One possibility, which corresponds to replica symmetry, is to assume that all (off-diagonal) elements of Q have the same value, say q. Then the interaction term in $\mathcal{H}_Q(s)$ is proportional to $\sum_{(ab)} s^a s^b$, which can be expressed in terms of $(\sum_a s^a)^2$. Using once again a dummy Gaussian integral,

one arrives at the expression

$$\begin{aligned} Z(Q) &= \exp\left(-\frac{nq}{2T^2}\right) \int \frac{\mathrm{d}z}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2} \sum_{s^1,\dots,s^n} \exp\left(\frac{z\sqrt{q}+h}{T} \sum_a s^a\right) \\ &= \exp\left(-\frac{nq}{2T^2}\right) \int \frac{\mathrm{d}z}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2}\, \exp\left\{n \ln\left[2\cosh\left(\frac{z\sqrt{q}+h}{T}\right)\right]\right\}, \end{aligned} \tag{11.172}$$

because the sum of the exponential splits into n independent factors. On the other hand, since $Q_{ab} = q$, $\forall a \neq b$, one has $\sum_{(ab)} Q_{ab}^2 = n(n-1)q^2/2$. We can now substitute in eq. (11.170), obtaining

$$\begin{aligned} A(q) = &-\frac{n}{4T^2} + \frac{nq}{2T^2} + n(n-1)\frac{q^2}{4T^2} \\ &- \ln\left\{\int \frac{\mathrm{d}z}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2}\, \exp\left[n \ln\left[2\cosh\left(\frac{z\sqrt{q}+h}{T}\right)\right]\right]\right\}. \end{aligned} \tag{11.173}$$

We can now take the $n \to 0$ limit, obtaining

$$A(q) = -\frac{(1-q)^2}{4T^2} - \int \frac{\mathrm{d}z}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2} \ln\left[2\cosh\left(\frac{z\sqrt{q}+h}{T}\right)\right]. \tag{11.174}$$

By imposing $\partial A/\partial q = 0$, and integrating by parts, we obtain the saddle-point equation

$$q = \int \frac{\mathrm{d}z}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2} \tanh^2\left(\frac{z\sqrt{q}+h}{T}\right). \tag{11.175}$$

We now set $h = 0$. By expanding around $q = 0$ the equation takes the form

$$q = \int \frac{\mathrm{d}z}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2} \left(z^2\frac{q}{T^2} - \frac{2}{3}z^4\frac{q^2}{T^4} + \cdots\right) = \frac{q}{T^2} - 2\frac{q^2}{T^4} + \cdots. \tag{11.176}$$

Thus $q = 0$ for $T > T_c = 1$, while there is a positive solution for $T < T_c$. We would like to accept this as the signature of a phase transition for $T = T_c$, and to take q as the corresponding order parameter.

Unfortunately, this solution has problems. Let us look at its properties at small T. Since, for $T \ll 1$, one has

$$\tanh^2\left(\frac{z\sqrt{q}}{T}\right) \simeq 1 - 4\exp\left(-2\,|z|\,\frac{\sqrt{q}}{T}\right), \tag{11.177}$$

we obtain the approximate equation

$$q \simeq 1 - 4\int_{-\infty}^{\infty} \frac{\mathrm{d}z}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2} \exp\left(-2\,|z|\,\frac{\sqrt{q}}{T}\right) \simeq 1 - \sqrt{\frac{2}{\pi q}}\,T. \tag{11.178}$$

This has the solution

$$q \simeq 1 - \sqrt{\frac{2}{\pi}}\,T, \tag{11.179}$$

which yields

$$A(q) \simeq -\frac{(1-q)^2}{4T^2} - \int \frac{\mathrm{d}z}{\sqrt{2\pi}}\, \mathrm{e}^{-z^2/2} \frac{|z|\sqrt{q}}{T} \simeq -\frac{1}{2\pi} - \frac{1}{T}\sqrt{\frac{2}{\pi}}. \tag{11.180}$$

Now the free energy per spin is given by $f = -T\,A(q)$, and the entropy by $s = -\partial f/\partial T$, We conclude that for $T \to 0$ we would have

$$s(0) = -\frac{1}{2\pi}. \tag{11.181}$$

Thus the entropy at $T = 0$ is negative. Since the entropy is the logarithm of the number of accessible configurations, this result is clearly nonphysical.

The origin of the problem was pointed out by de Almeida and Thouless [Alme78], who noticed that the replica-symmetric solution is unstable in a region of the (T, h) plane that contains the whole line $h = 0$, $T < T_c$. The instability is exhibited by the presence of negative eigenvalues of the Hessian $\partial^2 A/\partial Q^{ab}\,\partial Q^{cd}$. Now, negative eigenvalues would yield imaginary results for the saddle-point integration, which involves the square root of the determinant of the Hessian.

The solution proposed by Parisi [Pari79] was to break the replica symmetry into an iterative scheme. One forms n/m_1 groups of m_1 replicas, and assigns a value q_1 to Q_{ab} if a and b belong to the same group, and lets them q_0 otherwise. One then forms m_1/m_2 groups of replicas within each of the n/m_1 groups, and assigns to them the value q_2 if they belong to the same group, and lets them remain q_1 otherwise. The same procedure is iterated at the level of the m_1/m_2 groups, and so on. At the kth step of replica symmetry breaking, one has a Q matrix parametrized by $(k+1)$ values of q, $(q_0, q_1, \ldots, q_k)$, and k values of m, $1 < m_1 < m_2 < \cdots < m_k < n$. Now one takes the $n \to 0$ limit. The parameters m_ℓ were defined in the interval $1 \le m \le n$, but in the $n \to 0$ limit, the m become continuous parameters in the interval $0 \le m \le 1$. One can then optimize with respect to the qs and the ms. As a result, one obtains a function $q(x)$ defined on the interval $0 \le x \le 1$, such that $q(x) = q_\ell$ if $m_{\ell-1} < x \le m_\ell$. In the limit $k \to \infty$, this converges to a continuous, monotonically increasing function $q(x)$ defined on the same interval. Its inverse, $x(q)$, can be interpreted as the cumulative distribution of the overlap q, that is, the probability that two identical but independent replicas of the system are found in states whose average reciprocal overlap does not exceed q.

What is the interpretation of this result? It is easier to visualize it at a finite level, say k, of replica symmetry breaking. As we have seen, the value of Q_{ab} represents the average overlap between the configurations of replica a and replica b. This overlap is largest and equal to q_k if the two replicas belong to the same group at the kth level of symmetry breaking. If they instead belong to different groups that are subgroups of, say, the $(k-1)$th level, their mutual overlap equals q_{k-1}. This means that the system can be found in a large number (in the limit, infinite) of *pure states*. In each of these states the configurations are rather close to one another, since the overlap $q_k > 0$ can be interpreted as the Edwards-Anderson order parameter $\overline{\langle\sigma_i\rangle}^2$. However, the states are hierarchically arranged, such that, given three such states, say, α, β, and γ, one has, e.g., $q_{\alpha\beta} = q_{\alpha\gamma} \le q_{\beta\gamma}$. Interpreting $1/q$ as a measure of distance, this means that in this situation all triangles are isosceles, with the longest sides

equal in length, a property called **ultrametricity**, which describes a structure in which the pure states are organized hierarchically (see, for example, the review [Ramm86]). These properties have been investigated by rigorous analysis, in particular by S. Ghirlanda and F. Guerra [Ghir98].

The replica-trick calculation is clearly heuristic, but the results have been confirmed rigorously by other methods, in particular by F. Guerra [Guer02, Guer03] and M. Talagrand [Tala06]. A short review is found in [Pari09], while a comprehensive treatment is due to M. Talagrand [Tala11a, Tala11b]. Unfortunately, I cannot discuss here these alternative approaches. One can find an accessible introduction to one of such methods, the cavity method, in [Meza15].

11.6 The Hopfield Model of Associative Memory

> Why are we using spin glasses to understand the brain?
> Because we aren't able to use the brain to understand
> spin glasses.
>
> —M. A. Virasoro

The spin glass has been used as a paradigm for complex systems of a very different nature. For example, one can represent elements of the brain cortex as a network of formal neurons in interaction. A **formal neuron** is a two-state automaton (introduced by McCulloch and Pitts in 1943 [McCu43]) that represents the dynamic behavior of a neural cell (neuron) in a very simplified manner. In the brain, we expect that the interaction between a given pair of neurons is in general different from that between a different pair, so that the network can be represented as an assembly of formal neurons with random interactions, similarly to a spin glass. In the present section, after giving a very cursory introduction to the physiology of neurons and their connections, I discuss an idealized dynamical model of a network of formal neurons, and show by statistical mechanics methods that such a system can act as a content-adressable memory. This model was introduced by J. J. Hopfield [Hopf82], and is known therefore as the **Hopfield model**.

The fundamental structure of a neuron from the cortex is shown in fig. 11.10. Two series of tree-like projections depart from the cell body—on the one hand, the **dendrites**, on the other, the **axon** (a more or less cylindrical projection that can reach respectable lengths), which in its turn branches out into a large number of terminal branches. These are in direct contact with the dendrites of other nerve cells. The contacts are called **synapses** and are endowed with a complex structure.

The activity of a neuron is manifested in the emission of an **action potential**—an electrochemical excitation of invariable amplitude that is born in the cell body and moves at constant speed along the axon. When the action potential reaches the terminal branches, it excites the synapses, which discharge special chemical substances (the **neurotransmitters**) in the intercellular medium.

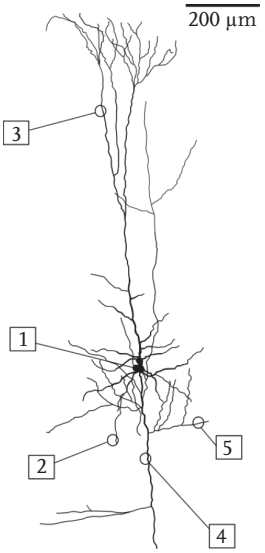


FIGURE 11.10. Scheme of a pyramidal neuron in the cortex. (1) Soma, (2) basal dendrite, (3) apical dendrite, (4) axon, (5) collateral axon. Adapted from *Wikipedia*. Image by Fabuio, licensed under the Creative Commons Attribution 4.0 International license.

The neurotransmitters are recognized by particular receptors placed on the cell membrane *downstream* from the synapse. The emission of the action potential is determined, in a complex and as yet not completely known fashion, by the activity of the receptors. McCulloch and Pitts suggested representing each neuron state by means of a binary variable $\tau : \tau = 0$ if the neuron is inactive (does not emit action potential), $\tau = 1$ if it is active (is emitting action potential, or is in the *refractory period* that follows each emission, in which it cannot emit the action potential again).

The probability $p_i(t)$ that a given neuron i will be active at the next instant t is therefore given by an expression containing the activities τ_j of all neurons upstream from neuron i:

$$p_i(t) = f\left(\sum_j J_{ij}\tau_j - \theta_i\right). \tag{11.182}$$

In this expression, the coefficients J_{ij} are defined for each ordered pair (i, j) of neurons and represent the effectiveness of the synaptic connection between neuron j (upstream) and neuron i (downstream). The function $f(x)$ is monotonically increasing. Then a positive value of J_{ij} implies that the synapse is an **exciter synapse**—in other words, that it tends to excite the downstream neuron when the one upstream is excited. A negative value of J_{ij} implies that the synapse is an **inhibitor synapse**—in other words, that it tends to suppress the excitation of the downstream neuron when the upstream neuron is excited. The quantity θ_i (which varies from neuron to neuron) is a threshold that determines the dendrite excitation level beyond which it is probable that the downstream neuron becomes active.

The function $f(x)$ is not determined—for simplicity, one usually assumes that it is the same for all neurons. The simplest case is obviously a simple step function:

$$f(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1, & \text{if } x > 0. \end{cases} \tag{11.183}$$

Obviously, however, Fermi-like functions, such as

$$f(x) = \frac{1}{1 + e^{-\beta x}}, \tag{11.184}$$

are also acceptable. In this case, the parameter β plays to some extent the role of the inverse of temperature.

To define a model for this system, one usually assumes that a synchronization mechanism between neurons does not exist. Then the evolution of the system can be described as follows: during each time interval, a neuron i is chosen (at random) and the expression $h_i = \sum_j J_{ij}\tau_j - \theta_i$, is evaluated; one then sets $\tau_i = 1$ with a probability equal to $f(h_i)$. This mechanism is known as **asynchronous updating**.

If we assume that J_{ij} is a symmetrical matrix and that $f(x)$ is given by eq. (11.183), it is easy to see that the dynamical process so defined coincides with the Monte Carlo dynamics relative to an Ising system, whose variables are represented by the τ_is, and whose Hamiltonian is equal to

$$\mathcal{H}(\tau) = -\frac{1}{2}\sum_{i \neq j} J_{ij}\tau_i\tau_j + \sum_i \theta_i\tau_i. \tag{11.185}$$

The transformation $\tau_i \to \sigma_i$, where

$$\tau_i = \frac{1 + \sigma_i}{2} \tag{11.186}$$

transforms this Hamiltonian into that of the usual Ising model, in the variables $\sigma_i = \pm 1$. The dynamics looks like a zero-temperature Monte Carlo simulation, in which each chosen spin flips to the state of lower energy. In this particular case, therefore, we can expect the dynamically stable states of the neural network to be the local minima of the Hamiltonian. We can then convince ourselves that by taking expression (11.184) instead, one obtains the Monte Carlo dynamics of the same system at a temperature given by $k_B T = \beta^{-1}$. In this case, we expect that the long-term behavior will be described by the statistical mechanics of this system—more particularly, by the thermodynamic equilibrium phases corresponding to the Hamiltonian (11.185).

Hopfield's crucial observation [Hopf82] is that a system of this type behaves as an **associative memory**, which is addressable by content. In standard computer memory, it is necessary to identify the address of each stored item in order to retrieve its contents. In associative memory, the stored items can be retrieved if one inputs the system with partial information about them. It is possible, for instance, to retrieve a name like *Hopfield* starting from a substring (with errors) like *Kopfi*.

In Hopfield's model, the Hamiltonian of the system is given by

$$\mathcal{H}(\sigma) = -\frac{1}{2N}\sum_{ij} J_{ij}\sigma_i\sigma_j, \tag{11.187}$$

where the interaction matrix $\mathsf{J} = (J_{ij})$ is given by

$$J_{ij} = \sum_{\alpha=1}^{p} \xi_i^\alpha \xi_j^\alpha. \tag{11.188}$$

Here the **patterns** $\xi^\alpha = (\xi_i^\alpha)$ are p binary vectors of length N that represent the information we wish to encode in the network. One assumes that the patterns have no special structure, that is, that they can be represented by sequences of independent variables assuming the values ± 1. As a consequence, the overlap between different patterns is of order $1/\sqrt{N}$.

This model has been solved by Amit, Gutfreund, and Sompolinsky [Amit85]. If p is finite, the partition function of the system is given by

$$\begin{aligned} Z &= \sum_\sigma \exp\left(-\frac{p}{2k_\mathrm{B}T}\right) \exp\left[\frac{1}{2Nk_\mathrm{B}T}\sum_\alpha\left(\sum_i \xi_i^\alpha \sigma_i\right)^2\right] \\ &= \left(\frac{N}{k_\mathrm{B}T}\right)^{p/2} \mathrm{e}^{-p/2k_\mathrm{B}T} \int \prod_\alpha \frac{\mathrm{d}m^\alpha}{\sqrt{2\pi}} \exp\left\{-\frac{Nm^2}{2} + \sum_i \ln\left[2\cosh\left(\frac{\boldsymbol{m}\cdot\boldsymbol{\xi}_i}{k_\mathrm{B}T}\right)\right]\right\}. \end{aligned} \tag{11.189}$$

Here we have introduced the vectors $\boldsymbol{\xi}_i = (\xi_i^1, \ldots, \xi_i^p)$ for each i, and the dummy vector $\boldsymbol{m} = (m^1, \ldots, m^p)$. As long as p is finite, the integral is dominated by its saddle-point value, yielding

$$f(T) = -\frac{k_\mathrm{B}T}{N}\ln Z = \frac{1}{2}m^2 - \frac{k_\mathrm{B}T}{N}\sum_i \ln\left[2\cosh\left(\frac{\boldsymbol{m}\cdot\boldsymbol{\xi}_i}{k_\mathrm{B}T}\right)\right]. \tag{11.190}$$

The order parameter $\boldsymbol{m}$ is determined by the saddle-point equation $\partial \ln Z/\partial \boldsymbol{m} = 0$, corresponding to

$$\boldsymbol{m} = \frac{1}{N}\sum_i \boldsymbol{\xi}_i \tanh\left(\frac{\boldsymbol{m}\cdot\boldsymbol{\xi}_i}{k_\mathrm{B}T}\right) = \overline{\boldsymbol{\xi}\tanh\left(\frac{\boldsymbol{m}\cdot\boldsymbol{\xi}}{k_\mathrm{B}T}\right)}. \tag{11.191}$$

Since the $\boldsymbol{\xi}_i$ are independent, in the $N\to\infty$ limit, expressions of the form $(1/N)\sum_i$ converge to their averages by the law of large numbers. (This happens only as long as p is finite.) We thus obtain the mean-field equations where the overline denotes average with respect to the distribution of the $\boldsymbol{\xi}$.

Assume that the vector $\boldsymbol{m}$ has only one nonvanishing component, say, $m^1 = m$. We then obtain

$$m = \tanh\left(\frac{m}{k_\mathrm{B}T}\right), \tag{11.192}$$

which is the self-consistency equation for the mean-field Ising model, and which has nonvanishing solutions for $k_\mathrm{B}T < k_\mathrm{B}T_\mathrm{c} = 1$. Thus, for temperatures below T_c, the system can be found in one of several equilibrium states, each having a finite overlap with one, and only one, of the stored patterns. The overlap goes to 1 as T goes to zero. It turns out that solutions of eq. (11.191) with more than one nonvanishing component of m have a larger free energy, and correspond therefore to metastable states.

Given such a system, prepare it in a state that has a finite, even if small, overlap with one of the stored patterns, and let the system evolve at a sufficiently low temperature. (Actually it may be useful to start at a comparatively high temperature, but always below the mean-field critical temperature T_c.) The system then evolves following the asynchronous updating rule we discussed and that we have seen is equivalent to a Monte Carlo dynamics.

After a while, the system will be in one of the equilibrium phases of the model that has a finite overlap with the chosen stored pattern. Then, slowly reducing the temperature to zero allows one to retrieve the entire ("infinite") pattern with a small number of errors. In this way, the evolution of Hopfield's model allows one to retrieve the information that was not completely present in the stimulus but was stored in the interaction coefficients.

The Hopfield model is one of the simplest cases of application of the concepts of statistical physics to the description of a biological system.

Exercise 11.22 Show that if $k_B T < 1$ there exist solutions to eqs. (11.191) of the form

$$\boldsymbol{m} = (m^i), \qquad m^i = m, \quad i = 1, \dots, r; \qquad m^i = 0, \quad i = r+1, \dots, p.$$

Show that these solutions are metastable, since the corresponding free energy is larger than that of the solution with $r = 1$.

Exercise 11.23 (Capacity of a Hopfield network) The maximal number $p_{\max}$ of patterns that can be stored error-free in a Hopfield network can be evaluated as follows. Define the Hopfield Hamiltonian $\mathcal{H}(\sigma; \xi)$ by the following (slightly modified) expression

$$\mathcal{H}(\sigma; \xi) = -\sum_{\alpha=1}^{p} \left(\sum_{i=1}^{N} \xi_i^\alpha \sigma_i \right)^n , \tag{11.193}$$

where $n = 2$ and σ and the patterns ξ^α are vectors of length N with entries equal to ± 1. Assume that the entries ξ_i^α are independent and identically distributed, equal to $+1$ with probability $\frac{1}{2}$.

1. Give the expression of the energy difference $\Delta E = \mathcal{H}(\hat{\xi}^1; \xi) - \mathcal{H}(\xi^1; \xi)$, where $\hat{\xi}^1$ differs from the stored pattern ξ^1 by a single spin flip: $\hat{\xi}_i^1 - \xi_i^1 = -2\xi_i^1 \delta_{i i_0}$. Evaluate the average $\langle \Delta E \rangle$ and variance $\Sigma^2 = \operatorname{Var} \Delta E$ of ΔE as functions of p and N, provided p and N are large.

2. The i_0th bit becomes unstable when the magnitude of the fluctuation exceeds the energy gap $\langle \Delta E \rangle$ and the sign of the fluctuation is opposite of the sign of the energy gap. Evaluate the probability that this takes place.

3. Deduce the maximal value of p such that this probability does not exceed a small threshold (say, of 0.5%).

4. Repeat the evaluation for $n > 2$. How does the capacity depend on n?

Recommended Reading

The theory of linear polymers is presented in a simple fashion in M. Doi, *Introduction to Polymer Physics*, Oxford, UK: Clarendon Press, 1996. One can also usefully consult P.-G. de Gennes, *Scaling Concepts in Polymer Physics*, Ithaca, NY: Cornell University Press, 1979. A discussion of percolation theory is due to D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2nd ed., London: Taylor and Francis, 1992. An introduction to

spin-glass theory is provided by K. H. Fischer and J. A. Hertz, *Spin Glasses*, Cambridge, UK: Cambridge University Press, 1991. An important collection of articles on spin glasses is contained in M. Mézard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond*, Singapore: World Scientific, 1987. A comprehensive introduction to the replica theory is V. Dotsenko, *Introduction to the Replica Theory of Disordered Statistical Systems*, Cambridge, UK: Cambridge University Press, 2001. The theory of spin glasses has recently found several unexpected applications: a very didactic presentation can be found in O. C Martin, R. Monasson, and R. Zecchina, "Statistical mechanics methods and phase transitions in optimization problems," *Theoretical Computer Science* 265, 3–67 (2001), available at arXiv:cond-mat/0104428. One can also read with benefit A. Hartmann and M. Weigt, *Phase Transitions in Combinatorial Optimization Problems*, Weinheim: Wiley-VCH, 2005. A thorough exposition of these applications is contained in M. Mézard and A. Montanari, *Information, Physics, and Computation*, Oxford: Oxford University Press, 2009. The developments of the theory of neural networks based on the Hopfield approach are discussed in D. J. Amit, *Modelling Brain Function: The World of Attractor Neural Networks*, Cambridge, UK: Cambridge University Press, 1989.

Appendix A Probability Refresher

A.1 Events and Probability

Let us consider a set Ω of **elementary events** (also called a **sample space**) such that only one of them must in any case occur as the result of a probabilistic phenomenon. The *measurable* subsets of Ω will be called **events**. Let us, for instance, consider the throw of a die. The set of elementary events is $\Omega = \{1, 2, 3, 4, 5, 6\}$; some of the possible events (subsets of Ω) are the *even number* event $\{2, 4, 6\}$, the *greater than or equal to 4* event $\{4, 5, 6\}$, the *greater than 6* event $\{\} = \emptyset$, and so on.

A **probability** $p(\alpha)$ is a real number associated with each event $\alpha \subseteq \Omega$ that satisfies the following requirements:

- $0 \leq p(\alpha) \leq 1, \forall \alpha \subseteq \Omega$;
- $p(\emptyset) = 0$ and $p(\Omega) = 1$;
- $p(\alpha \cup \beta) = p(\alpha) + p(\beta) - p(\alpha \cap \beta), \forall \alpha, \beta \subseteq \Omega$.

More specifically, if α and β are **mutually exclusive** events, that is, if $\alpha \cap \beta = \emptyset$, then one has $p(\alpha \cup \beta) = p(\alpha) + p(\beta)$. Then, given a set of mutually exclusive events $\{\alpha_i\}$ such that $\bigcup_i \alpha_i = \Omega$, the probability of a generic event β is given by

$$p(\beta) = \sum_i p(\beta \cap \alpha_i). \tag{A.1}$$

This is known as the **law of total probabilities**. With a judicious choice of events $\{\alpha_i\}$, this result can simplify the calculation of the probabilities of complex events, reducing it to the calculation of the probabilities of simpler events.

It is customary to denote an event whose probability equals 1 as **almost certain** (and one says that it happens **almost surely**), and one whose probability vanishes as **almost impossible** (or that it happens **almost never**). Indeed, there might no intrinsic impossibility for an almost impossible event to take place (for instance, that a coin lands on head an infinite number of times), but this event has a vanishing probability.

A.2 Random Variables

A **random variable** is a real function defined on Ω, that is, a variable whose values depend on the outcome of a probabilistic phenomenon. The random variable can be **discrete**, when its values belong to a discrete set, or **continuous**, when it can in principle assume any real value in a given interval. A random variable can be characterized by events and by associated probabilities. For a discrete random variable x, the probability p_i of the event that it takes the value i defines its **probability distribution** (sometimes called *probability mass function*). Otherwise, if x is a continuous random variable, one can consider events such as $a \leq x < b$, where a and $b \geq a$ are two real numbers. In many cases, the probabilities of these events are *differentiable functions* of the extremes (a, b) of the interval. One can then define the **probability distribution** $P(x)$ (sometimes called the *probability density function*) of the random variable x, such that

$$p(a \leq x < b) = \int_a^b \mathrm{d}x\, P(x). \tag{A.2}$$

Some simple examples are the following.

The binomial distribution: It is the distribution of the number of successes in N replicas of the same experiment, where all replicas are independent (that is, the probability of the outcome of each replica does not depend on the outcome of any other) and the success probability equals p for each experiment. It is a discrete distribution, given by

$$p_n = \binom{N}{n} p^n (1-p)^{N-n}, \qquad 0 \leq n \leq N. \tag{A.3}$$

The Poisson distribution: It is the limit of the binomial distribution for $N \to \infty$ with $p\,N = \mu$ kept constant. It is given by

$$p_n = \frac{\mathrm{e}^{-\mu}}{n!}\mu^n, \qquad 0 \leq n < +\infty. \tag{A.4}$$

The uniform distribution: As indicated by its name, it is the distribution of a continuous random variable assuming values in an interval $a \leq x \leq b$, where the probability of finding the variable in any subinterval contained in it depends only on the length of the subinterval. It is given by

$$P(x) = \begin{cases} 1/(b-a), & \text{if } a \leq x \leq b; \\ 0, & \text{otherwise.} \end{cases} \tag{A.5}$$

The Gaussian distribution: It is a distribution for a continuous variable that can assume any real value. It is given by

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-x_0)^2}{2\sigma^2}\right], \qquad -\infty < x < +\infty. \tag{A.6}$$

Its most probable value is equal to x_0, and the probability of a value x different from x_0 decreases very rapidly as $|x - x_0|$ increases beyond a value of order σ.

The delta distribution: This is a singular distribution, represented by the Dirac delta "function" $\delta(x - x_0)$, which satisfies

1. for any continuous function $F(x)$, one has $\int \mathrm{d}x\, \delta(x - x_0)\, F(x) = F(x_0)$;
2. $\delta(-x) = \delta(x)$;
3. $\delta(a\, x) = |a|^{-1}\, \delta(x)$.

The delta distribution $P(x) = \delta(x - x_0)$ corresponds to the situation where the random variable x assumes the value x_0 with certainty. For example, the distribution of the results of the throws of an unbiased die can be represented, taking advantage of the law of total probability, by

$$P(x) = \sum_{k=1}^{6} \frac{1}{6}\, \delta(x - k). \tag{A.7}$$

A.3 Averages and Moments

The average of a random variable x is defined by

$$\langle x \rangle = \begin{cases} \sum_x x p_x, & \text{if } x \text{ is discrete;} \\ \int \mathrm{d}x\, x\, P(x), & \text{if } x \text{ is continuous.} \end{cases} \tag{A.8}$$

A function $f(x)$ of a random variable x is itself a random variable. One has, for example,

$$\langle f \rangle = \int \mathrm{d}f\, f\, P(f) = \int \mathrm{d}x\, \frac{\mathrm{d}f}{\mathrm{d}x} f(x)\, P(f(x)). \tag{A.9}$$

Assume for a moment that $f(x)$ is monotonically increasing in x. Then if $f_0 < f < f + \mathrm{d}f$, one has $x_0 < x < x + \mathrm{d}x$, where $f(x_0) = f_0$ and $\mathrm{d}x = (f'(x_0))^{-1} \mathrm{d}f$. We obtain therefore $(\mathrm{d}f/\mathrm{d}x)\, P(f(x_0)) = P(x_0)$, leading to

$$\langle f \rangle = \int \mathrm{d}x f(x)\, P(x). \tag{A.10}$$

It is easy to convince oneself that this relation holds more generally. One thus obtains, for example, the useful relation

$$\langle \delta(x - x_0) \rangle = \int \mathrm{d}x\, \delta(x - x_0)\, P(x) = P(x_0). \tag{A.11}$$

The averages of the powers of a random variable are called **moments**. One therefore has

$$m_1(x) = \langle x \rangle = \int \mathrm{d}x\, x\, P(x); \tag{A.12}$$

$$m_2(x) = \langle x^2 \rangle = \int \mathrm{d}x\, x^2\, P(x); \tag{A.13}$$

$$\vdots$$

Table A.1 Average, second moment, and variance of the elementary probability distributions.

Distribution	$\langle x \rangle$	$\langle x^2 \rangle$	Var x
Binomial	pN	$p^2 N(N-1) + pN$	$Np(1-p)$
Poisson	μ	$\mu(\mu+1)$	μ
Uniform	$(a+b)/2$	$(a^2+ab+b^2)/3$	$(a-b)^2/12$
Gaussian	x_0	$x_0^2+\sigma^2$	σ^2
Delta	x_0	x_0^2	0

The **variance** of x is defined by

$$\operatorname{Var} x = \left\langle (x - \langle x \rangle)^2 \right\rangle = \left\langle x^2 \right\rangle - \langle x \rangle^2. \tag{A.14}$$

One obviously has $\operatorname{Var} x \geq 0$, and therefore the variance is often denoted by σ^2. The positive square root of the variance of x is called the **standard deviation**, is denoted by $\operatorname{std} x$ or σ, and provides a measure of the typical deviation of x from its average.

The averages, second moments, and variances of the elementary probability distributions are summarized in table A.1.

A.4 Conditional Probability: Independence

Given two events, α and β, the **conditional probability** of β, given α, is defined by

$$p(\beta|\alpha) = \frac{p(\alpha \cap \beta)}{p(\alpha)}. \tag{A.15}$$

In particular, if α is the certain event, that is, if $\alpha = \Omega$, it coincides with $p(\beta)$. The law of total probability can be then reformulated as follows: given a set $\{\alpha_i\}$ of mutually exclusive events, $\alpha_i \cap \alpha_j = \emptyset$ $(i \neq j)$, such that $\bigcup_i \alpha_i = \Omega$, the probability of a generic event β is given by

$$p(\beta) = \sum_i p(\beta|\alpha_i) p(\alpha_i). \tag{A.16}$$

Given two events α and β, the probability $p(\alpha \cap \beta)$ of their intersection is called their **joint probability**. Two events are call **independent** if $p(\alpha \cap \beta) = p(\alpha)p(\beta)$. For independent events, one therefore has

$$p(\beta|\alpha) = p(\beta); \qquad p(\alpha|\beta) = p(\alpha). \tag{A.17}$$

Given a pair (x, y) of continuous random variables, one can consider the corresponding events $\alpha = (a_1 \leq x < b_1)$ and $\beta = (a_2 \leq y < b_2)$. We can then define the **joint distribution** $P(x, y)$ analogously to the joint probability:

$$p(\alpha \cap \beta) = \int_{a_1}^{b_1} \mathrm{d}x \int_{a_2}^{b_2} \mathrm{d}y \, P(x, y). \tag{A.18}$$

If, for any values of the interval extremes, one obtains independent events of this kind, the variables x and y are called *independent*, and one has

$$P(x, y) = P_x(x)\, P_y(y). \tag{A.19}$$

Analogous relations hold for discrete random variables. It is easy to see that, for independent variables, one has

$$\langle xy \rangle = \langle x \rangle \langle y \rangle . \tag{A.20}$$

The inverse is not true, but *if and only if* x and y are independent, this relation holds for all moments

$$\langle x^p y^q \rangle = \langle x^p \rangle \langle y^q \rangle, \qquad \forall p, q \in \mathbb{N}. \tag{A.21}$$

If this is the case only for $p = q = 1$, the variables are called **uncorrelated**.

Let us now consider a certain number N of random variables $x_1, \ldots, x_N$. One has the following properties:

- The *average* of the sum of the random variables (*even if not independent*) is equal to the sum of their averages:
$$\left\langle \sum_{i=1}^{N} x_i \right\rangle = \sum_{i=1}^{N} \langle x_i \rangle . \tag{A.22}$$
- The *variance* of the sum of *uncorrelated* variables is equal to the sum of the variances. In particular, if $\langle x_i \rangle = 0$, $\forall i$, one has
$$\left\langle \left(\sum_{i=1}^{N} x_i \right)^2 \right\rangle = \sum_{i=1}^{N} \langle x_i \rangle^2 . \tag{A.23}$$
Independent variables are a fortiori uncorrelated, and therefore this relation holds for them.

The intuitive notion that independent, identically distributed (i.i.d.) random variables are usually close to their average is expressed by **Chebyshev's inequality**, which stipulates the probability that a random variable X with average μ and variance σ^2 differs from μ more than the arbitrary positive value Δ cannot exceed σ^2/Δ^2:

$$\text{Prob}\,(|X - \mu| \geq \Delta) \leq \frac{\sigma^2}{\Delta^2}. \tag{A.24}$$

This may be proven as follows. Denote by $\langle X | q(X) \rangle$ the **conditional expectation** of the random variable X subject to the condition $q(X)$, that is, the average of X over all values that satisfy the condition, divided by the probability that the condition is satisfied. We then have

$$\begin{aligned} \sigma^2 = \langle |X - \mu|^2 \rangle &= \left\langle |X - \mu|^2 \,\middle|\, \Delta \leq |X - \mu| \right\rangle \text{Prob}\,(\Delta \leq |X - \mu|) \\ &\quad + \left\langle |X - \mu|^2 \,\middle|\, \Delta > |X - \mu| \right\rangle \text{Prob}\,(\Delta > |X - \mu|) \\ &\geq \Delta^2 \ \text{Prob}\,(\Delta \leq |X - \mu|) + 0 \cdot \text{Prob}\,(\Delta > |X - \mu|) \\ &= \Delta^2 \ \text{Prob}\,(\Delta \leq |X - \mu|) . \end{aligned} \tag{A.25}$$

The result follows by dividing by Δ^2. Thus, for example, the probability of finding X further away than 2σ from μ is smaller than 25%.

A.5 Characteristic Functions

The **characteristic function** of a random variable x is defined by

$$\Gamma(k) = \langle \exp(\mathrm{i}kx) \rangle = \int \mathrm{d}x \, \exp(\mathrm{i}kx) \, P(x). \tag{A.26}$$

The characteristic function is related to the moments $\langle x^\ell \rangle$ ($\ell = 1, 2, \ldots$) by

$$\langle x^\ell \rangle = (-\mathrm{i})^\ell \left. \frac{\partial^\ell \Gamma}{\partial k^\ell} \right|_{k=0}. \tag{A.27}$$

If $\Gamma(k)$ is analytic around $k = 0$, all moments $\langle x^\ell \rangle$ exist, and one has

$$\Gamma(k) = \sum_{p=0}^{\infty} \frac{\mathrm{i}^p \, \langle x^p \rangle}{p!}. \tag{A.28}$$

The characteristic function of the sum of two *independent* random variables is the product of their respective characteristic functions:

$$\begin{aligned} \langle \exp[\mathrm{i}k(x+y)] \rangle &= \int \mathrm{d}x \int \mathrm{d}y \, \exp[\mathrm{i}k(x+y)] \, P(x) \, P(y) \\ &= \left(\int \mathrm{d}x \, \exp(\mathrm{i}kx) \, P(x) \right) \left(\int \mathrm{d}y \, \exp(\mathrm{i}ky) \, P(y) \right) \\ &= \langle \exp(\mathrm{i}kx) \rangle \langle \exp(\mathrm{i}ky) \rangle . \end{aligned} \tag{A.29}$$

This allows one to calculate the generating function of arbitrary linear combinations of independent variables. It is important to observe that the generating function of a Gaussian variable is itself a Gaussian:

$$\begin{aligned} \langle \exp(\mathrm{i}kx) \rangle &= \int \mathrm{d}x \, \exp(\mathrm{i}kx) \, \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-x_0)^2}{2\sigma^2} \right] \\ &= \exp\left(\mathrm{i}kx_0 - \frac{1}{2}k^2\sigma^2 \right). \end{aligned} \tag{A.30}$$

Therefore, linear combinations of independent Gaussian variables are also Gaussian, since the product of their Gaussian generating functions is a Gaussian.

It is straightforward to check that for small values of $|k|$, one has

$$\Gamma(k) \simeq \exp\left[\mathrm{i}k \, \langle x \rangle - \frac{1}{2}k^2\sigma^2 + \cdots \right], \tag{A.31}$$

where $\langle x \rangle$ is the average and σ^2 the variance of the random variable x, and the dots represent higher-order terms in k. We have indeed $\Gamma(0) = 1$ by the normalization condition, and, taking advantage of eq. (A.28),

$$\begin{aligned} \left.\frac{\partial \ln \Gamma}{\partial k}\right|_{k=0} &= \mathrm{i}\,\langle x \rangle\,; \\ \left.\frac{\partial^2 \ln \Gamma}{\partial k^2}\right|_{k=0} &= \left.\frac{\partial}{\partial k}\left(\frac{1}{\Gamma(x)}\frac{\partial \Gamma}{\partial k}\right)\right|_{k=0} = \left.\frac{\partial^2 \Gamma}{\partial k^2}\right|_{k=0} - \left(\left.\frac{\partial \Gamma}{\partial k}\right|_{k=0}\right)^2 \\ &= -\langle x^2 \rangle + \langle x \rangle^2 = -\sigma^2. \end{aligned} \tag{A.32}$$

Therefore, by Taylor's theorem, we have $\ln \Gamma(k) = \mathrm{i}k\,\langle x \rangle - \sigma^2 k^2/2 + \mathrm{o}(k^2)$, and the result (A.31) follows. Here the notation $\mathrm{o}(x)$ means that

$$\lim_{x \to 0} \frac{\mathrm{o}(x)}{x} = 0. \tag{A.33}$$

Since $\left|\mathrm{e}^{\mathrm{i}kx}\right| = 1$ for all k and x real, the integral defining the characteristic function is absolutely convergent for all real values of k. However, this does not guarantee the analyticity of $\Gamma(k)$ and therefore the existence of all moments of $P(x)$. Consider, for example, the Cauchy distribution, defined by

$$P(x) = \frac{1}{\pi}\frac{\xi}{x^2 + \xi^2}, \qquad \xi > 0. \tag{A.34}$$

The integral defining $\langle x \rangle$ does not converge absolutely for this distribution, and therefore its value depends on how the integration limits are sent to infinity. Thus the first moment, and even more so the higher ones, does not exist. This is borne out by the expression of $\Gamma(k)$ that is obtained by a simple contour integral. The integrand has poles for $k = \pm \mathrm{i}\xi$ with residues $\pm \mathrm{e}^{\mp k\xi}/(2\pi)$ respectively, and the contour must be closed in the upper complex half-plane if $k > 0$, and in the lower one if $k < 0$. We obtain therefore

$$\Gamma(k) = \mathrm{e}^{-|k|\xi}, \tag{A.35}$$

which exhibits a singularity for $k = 0$.

A.6 Central Limit Theorem

Let $x_1, \ldots, x_N$ be N independent random variables, identically distributed, with average $\langle x \rangle$ and variance equal to σ^2. The **mean**, as opposed to the average, of these variables is a *random variable* defined by

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{A.36}$$

The **central limit theorem** states that, in the limit $N \to \infty$, the distribution of the mean tends to a Gaussian of average $\langle x \rangle$ and variance equal to σ^2/N. For the proof, it is sufficient

to consider the characteristic functions. Taking advantage of eq. (A.31), we have

$$\begin{aligned}\left\langle \exp\left(\mathrm{i}k\overline{x}\right)\right\rangle &= \left\langle \exp\left[\mathrm{i}k\left(\frac{1}{N}\sum_{i=1}^{N}x_i\right)\right]\right\rangle = \left\langle \exp\left(\frac{\mathrm{i}kx}{N}\right)\right\rangle^N \\ &= \left[\Gamma\left(\frac{k}{N}\right)\right]^N = \left[\exp\left(\frac{\mathrm{i}k\langle x\rangle}{N} - \frac{k^2\sigma^2}{2N^2} + \cdots\right)\right]^N \\ &= \exp\left[\mathrm{i}k\langle x\rangle - \frac{k^2\sigma^2}{2N} + \mathrm{o}(N^{-1})\right].\end{aligned} \tag{A.37}$$

Thus, if N is very large, the distribution is a Gaussian, very sharply peaked around $\langle x\rangle$, which means that the variable $\overline{x}$ assumes with probability close to 1 ("almost certainly") values very close to $\langle x\rangle$. This is known as the **law of large numbers**.

The theorem can be generalized to nonidentically distributed variables as long as no variable *dominates* the others (in other words, exhibits a variance that is much larger than that of the others)—it is necessary that the variance of each variable be finite. In that case, the average of the mean $\overline{x}$ approaches the average of the averages $\langle x_i\rangle$ and its variance approaches $1/N$ times the average of the variances.

For variables that are not strictly independent, the theorem remains true if, in the sequence, each variable is correlated with a finite number of neighboring variables.

A.7 Correlations

Given two random variables A and B, their **correlation** is given by

$$\mathrm{Corr}(A, B) = \langle A\,B\rangle - \langle A\rangle\,\langle B\rangle = \langle (A - \langle A\rangle)\,(B - \langle B\rangle)\rangle. \tag{A.38}$$

It obviously vanishes if A and B are independent. If it vanishes, A and B are said to be uncorrelated, but they are not necessarily independent.

We have, moreover, the **Cauchy-Schwarz** inequality

$$\left|\mathrm{Corr}(A, B)\right| \leq \sqrt{\mathrm{Var}\,A}\sqrt{\mathrm{Var}\,B}. \tag{A.39}$$

To obtain it, consider without loss of generality the case where $\langle A\rangle = \langle B\rangle = 0$. Given an arbitrary real number ξ, we have

$$0 \leq \langle (A - \xi\,B)^2\rangle = \langle A^2\rangle - 2\xi\,\langle AB\rangle + \xi^2\,\langle B\rangle^2\,. \tag{A.40}$$

Therefore the discriminant of the second-degree polynomial in ξ on the right-hand side must not be positive. This implies

$$\langle AB\rangle^2 \leq \langle A^2\rangle\langle B^2\rangle, \tag{A.41}$$

as we wanted to prove.

Let us consider a collection of N random variables $x_1, \ldots, x_N$. The **correlation matrix** $\mathsf{C} = (C_{ij})$ is defined by

$$C_{ij} = \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle, \qquad i, j = 1, \ldots, N. \tag{A.42}$$

It is symmetric and positive semidefinite. We have indeed, for an arbitrary vector $\xi = (\xi_i)$,

$$\sum_{ij} \xi_i \xi_j C_{ij} = \left\langle \left[\sum_i \xi_i (x_i - \langle x_i \rangle) \right]^2 \right\rangle \geq 0. \tag{A.43}$$

A collection of N variables is called (multivariate) **Gaussian** if their joint probability distribution has the form

$$P(x_1, \ldots, x_N) \propto \exp\left[-\frac{1}{2} \sum_{ij} A_{ij} (x - \langle x_i \rangle)(x_j - \langle x_j \rangle) \right]. \tag{A.44}$$

The matrix $\mathsf{A} = (A_{ij})$ must be symmetric and positive definite. The proportionality constant is defined by the **normalization condition**

$$\int \prod_{i=1}^{N} \mathrm{d}x_i \, P(x_1, \ldots, x_N) = 1, \tag{A.45}$$

and is equal to $(2\pi)^{-N/2} (\det \mathsf{A})^{1/2}$. To prove it, let us assume for simplicity that for each i, $\langle x_i \rangle = 0$. Define the new integration variables y_k by $y_k = \sum_j Q_{kj} x_j$, where the symmetrical matrix $\mathsf{Q} = (Q_{k\ell})$ satisfies $\mathsf{Q}^2 = \mathsf{A}$. The Jacobian of the transformation is given by $(\det \mathsf{Q})^{-1} = (\det \mathsf{A})^{-1/2}$. We thus obtain

$$\begin{aligned} &\int \prod_m \mathrm{d}x_m \, \exp\left(-\frac{1}{2} \sum_{k\ell} A_{k\ell} x_k x_\ell \right) \\ &= \int \prod_m \mathrm{d}y_m \, (\det \mathsf{A})^{-1/2} \, \exp\left(-\frac{1}{2} \sum_n y_n^2 \right) \\ &= (2\pi)^{N/2} (\det \mathsf{A})^{-1/2} . \end{aligned} \tag{A.46}$$

Let us prove that the correlation matrix $\mathsf{C} = (C_{ij})$ of a Gaussian collection of variables is equal to the inverse of the matrix A that defines the distribution: $\mathsf{C} = \mathsf{A}^{-1}$. By changing from the variables x to the variables y, the Jacobian cancels out with the $(\det \mathsf{A})^{1/2}$ of the normalization. The product $x_i x_j$ is now expressed by $\sum_k Q_{ik}^{-1} Q_{jk}^{-1} y_k^2$. We thus obtain

$$\begin{aligned} C_{ij} = \langle x_i x_j \rangle &= \int \prod_m \left(\frac{\mathrm{d}x_m}{\sqrt{2\pi}} \right) (\det \mathsf{A})^{1/2} \, x_i x_j \, \exp\left(-\frac{1}{2} \sum_{k\ell} A_{k\ell} x_k x_\ell \right) \\ &= \int \prod_m \left(\frac{\mathrm{d}y_m}{\sqrt{2\pi}} \right) \left(\sum_k Q_{ik}^{-1} Q_{jk}^{-1} y_k^2 \right) \exp\left(-\frac{1}{2} \sum_n y_n^2 \right) \\ &= \sum_k Q_{ik}^{-1} Q_{jk}^{-1} = A_{ij}^{-1}, \end{aligned} \tag{A.47}$$

as we wanted to prove.

A.8 Dirac's Delta "Function"

In this section we collect some useful properties of Dirac's delta "function" $\delta(x)$.

This "function" can be thought of as the limit of a peaked distribution, like the Gaussian one, as its width goes to zero. We then have, for any reasonably smooth function $f(x)$ defined on an interval containing x_0 in its interior,

$$\lim_{\sigma\to 0}\int_{-\infty}^{+\infty}\frac{\mathrm{d}x}{\sqrt{2\pi\,\sigma^2}}\,\mathrm{e}^{-(x-x_0)^2/2\sigma^2}f(x)=f(x_0). \tag{A.48}$$

We similarly have

$$\lim_{\sigma\to 0}\int_{-\infty}^{+\infty}\frac{\mathrm{d}x}{\pi}\,\frac{\sigma}{(x-x_0)^2+\sigma^2}f(x)=f(x_0). \tag{A.49}$$

We can interpret integrals containing Dirac's delta as a shorthand for limits of this kind, and write simply, by an abuse of notation,

$$\int \mathrm{d}x\,\delta(x-x_0)f(x)=f(x_0). \tag{A.50}$$

The reader is invited to prove the following properties of $\delta(x)$:

1. $\delta(-x)=\delta(x)$;
2. $\delta(\alpha x)=\delta(x)/\,|\alpha|$, $\forall\alpha\neq 0$;
3. $x\delta(x)=0$.

As a consequence of property 2., we have

$$\delta(f(x))=\sum_i\frac{\delta(x-x_i)}{|f'(x_i)|}, \tag{A.51}$$

where the x_i are the zeros of the reasonably smooth function $f(x)$.

It is also possible to define the derivative $\delta'(x)$ of the delta function, which satisfies

$$\int \mathrm{d}x\,\delta'(x)f(x)=-f'(0), \tag{A.52}$$

as can be checked by an integration by parts. Similar relations hold for higher derivatives.

We can define the delta "function" on a d-dimensional vector $\boldsymbol{r}=(r_1,\ldots,r_d)$ in the following way:

$$\delta(\boldsymbol{r})=\prod_{\alpha=1}^{d}\delta(r_\alpha). \tag{A.53}$$

We then have

$$\int \mathrm{d}\boldsymbol{r}\, \delta\,(\boldsymbol{r}-\boldsymbol{r}_0)\, f(\boldsymbol{r}) = f(\boldsymbol{r}_0), \tag{A.54}$$

where $f(\boldsymbol{r})$ is a smooth function defined in a region of $\mathbb{R}^d$ containing $\boldsymbol{r}_0$ in its interior. We then have, for example,

$$\int \mathrm{d}\boldsymbol{r}\, \nabla\delta\,(\boldsymbol{r}-\boldsymbol{r}_0)\, f(\boldsymbol{r}) = -\nabla f(\boldsymbol{r}_0). \tag{A.55}$$

Appendix B Convex Functions and the Legendre Transformation

B.1 Convex Functions

Many useful inequalities hold for **convex functions**. There is sometimes confusion about the definitions of concave and convex. The most used convention in mathematics is to define a $\cup$-shaped function as convex, and a $\cap$-shaped function as concave. A way to memorize the definition is to think that a $\cap$-shape looks like a cave, hence it is concave. More formally, a real-valued function $f(x)$ is **convex** over an interval (a, b) if for any $x_0, x_1 \in (a, b)$ and $0 \leq \alpha \leq 1$, one has

$$f\left((1-\alpha)x_0 + \alpha x_1\right) \leq (1-\alpha)f(x_0) + \alpha f(x_1). \tag{B.1}$$

See the scheme in fig. B.1. A function f is **concave** if $(-f)$ is convex. If the inequality is strict for any $x_0 \neq x_1$ and $0 < \alpha < 1$, f is said to be **strictly convex**.

If a function $f(x)$ has a second derivative which is continuous and nonnegative (positive) in (a, b), then f is convex (strictly convex). Indeed, given x_0 and x, by the Taylor expansion and by the weighted mean-value theorem there is a point u between x_0 and x such that

$$f(x) = f(x_0) + f'(x_0)\,(x - x_0) + \frac{1}{2}f''(u)\,(x - x_0)^2, \tag{B.2}$$

where the last term is nonnegative. Let us now consider $x_\alpha = (1-\alpha)x_0 + \alpha x_1$. We obtain

$$f(x_0) \geq f(x_\alpha) + f'(x_\alpha)\,(x_0 - x_\alpha) = f(x_\alpha) - f'(x_\alpha)\,\alpha\,(x_1 - x_0) \tag{B.3}$$

and

$$f(x_1) \geq f(x_\alpha) + f'(x_\alpha)\,(x_1 - x_\alpha) = f(x_\alpha) + f'(x_\alpha)\,(1-\alpha)\,(x_1 - x_0). \tag{B.4}$$

Multiplying (B.3) by $(1-\alpha)$ and (B.4) by α and summing, we obtain (B.1). The strict inequality is derived along the same lines when $f''(x) > 0$, $\forall x \in (x_0, x_1)$.

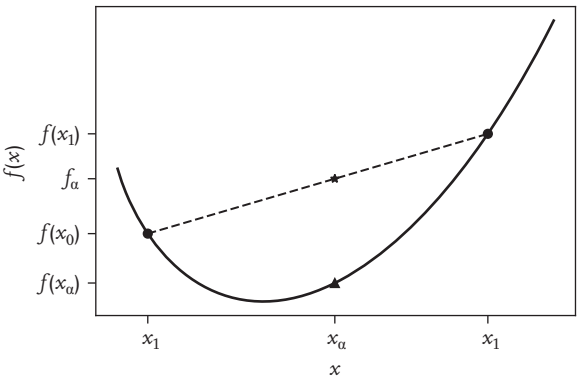


FIGURE B.1. Illustration of convexity. Given the point $x_\alpha = \alpha x_1 + (1-\alpha)x_0$, with $0 \le \alpha \le 1$, we have $f(x_\alpha) \le (1-\alpha)f(x_0) + \alpha f(x_1) = f_\alpha$, $\forall \alpha \in [0, 1]$.

We also have the converse result: if the convex function $f(x)$ is twice differentiable in the interval $[x_0, x_1]$, then $f''(x) \ge 0$ in the interval. Indeed, we have by definition, for any x for which $f''(x)$ exists,

$$f''(x) = \lim_{h \to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}. \tag{B.5}$$

On the other hand, by convexity, one has

$$\frac{1}{2}\Big(f(x+h) + f(x-h)\Big) \ge f(x). \tag{B.6}$$

Thus the quantity on the right-hand side of the relation above is non-negative for all $h \neq 0$. In the limit, we obtain $f''(x) \ge 0$. The strict inequality does not necessary hold, even if $f(x)$ is strictly convex. An example is the function $f(x) = x^4$.

Let us now take a convex function $f(x)$ that is twice differentiable in (a, b) and a point x_0 in the interior of the interval (a, b). Then $f(x) \ge f(x_0) + f'(x_0)(x - x_0)$ for $x \in (a, b)$. In other words, the graph of the convex function $f(x)$ lies above the tangent to it in x_0. The inequality is strict if $f(x)$ is strictly convex. Indeed, by eq. (B.2), we have, for some $u \in (x_0, x)$,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(u)(x - x_0)^2, \tag{B.7}$$

where the last term is nonnegative. Thus we have proved the weak inequality. We can prove the strict inequality by contradiction, assuming that $f(x)$ is strictly convex and that $f(x) = f(x_0) + f'(x_0)(x - x_0)$ for some $x \neq x_0$. Convexity then implies that, for any $u \in (x_0, x)$,

$$f(u) \le f(x)\frac{u - x_0}{x - x_0} + f(x_0)\frac{x - u}{x - x_0} = f'(x_0)(u - x_0) + f(x_0). \tag{B.8}$$

On the other hand, the result we have just derived implies

$$f(u) \ge f'(x_0)(u - x_0) + f(x_0), \tag{B.9}$$

and therefore

$$f(u) = f(x)\frac{u-x_0}{x-x_0} + f(x_0)\frac{x-u}{x-x_0}, \tag{B.10}$$

which violates the strict convexity.

B.2 The Jensen Inequality

The **Jensen inequality** states that if $f(x)$ is a convex function and x is a random variable, then

$$\langle f(x)\rangle \geq f(\langle x\rangle), \tag{B.11}$$

where $\langle x\rangle$ is the average of x over its distribution. If f is strictly convex, the equality in (B.11) implies that x is a constant.

To prove this result, we define $g(x) = f(\langle x\rangle) + f'(\langle x\rangle)(x - \langle x\rangle)$. Since $f(x)$ is convex, one has $f(x) \geq g(x)$, $\forall x$, with equality for $x = \langle x\rangle$. Moreover, if $f(x)$ is strictly convex, one has $f(x) > g(x)$, $\forall x \neq \langle x\rangle$. Since $g(x)$ is linear, $\langle g(x)\rangle = g(\langle x\rangle)$. Thus we have

$$\langle f\rangle \geq \langle g\rangle = g(\langle x\rangle) = f(\langle x\rangle). \tag{B.12}$$

The inequality is strict if $f(x)$ is strictly convex, unless x assumes only one value.

The Jensen inequality implies some useful corollaries. We often use it in cases in which $f(x)$ is an exponential function:

$$\langle \mathrm{e}^{-x}\rangle \geq \mathrm{e}^{-\langle x\rangle}, \tag{B.13}$$

for any real-valued random variable x. Taking the logarithm, this relation implies

$$\langle x\rangle \geq -\ln\langle \mathrm{e}^{-x}\rangle. \tag{B.14}$$

Let us now consider two vectors (p_x) and (q_x) with positive entries, with $p = \sum_x p_x$ and $q = \sum_x q_x$. The function $f(x) = x\ln x$ (for $x > 0$) is convex. Define

$$\mathfrak{S} = \sum_x p_x \ln\frac{p_x}{q_x}, \tag{B.15}$$

and the probability distribution

$$u_x = \frac{q_x}{q}. \tag{B.16}$$

We then have

$$\mathfrak{S} = q\sum_x u_x \left(\frac{p_x}{q_x}\right)\ln\frac{p_x}{q_x} = q\left\langle f\left(\frac{p_x}{q_x}\right)\right\rangle_u, \tag{B.17}$$

where $\langle \ldots \rangle_u$ is the average with respect to the probability distribution $u = (u_x)$. By the Jensen inequality, we obtain

$$\mathfrak{S} \geq qf\left(\left\langle \frac{p_x}{q_x} \right\rangle_u\right) = qf\left(\sum_x u_x \frac{p_x}{q_x}\right) = \left(\sum_x p_x\right) \ln \frac{\sum_{x'} p_{x'}}{\sum_{x'} q_{x'}} = p \ln \frac{p}{q}. \tag{B.18}$$

This relation is known as the **logsum inequality.**

B.3 Legendre Transformation

The Legendre transformation allows one, both in thermodynamics and in mechanics, to convert functions of a given set of variables (such as volume or velocity) to functions of their conjugate (such as pressure or momentum). Let us consider a function $f(x)$. For simplicity we assume that it is twice differentiable, and that its second derivative is positive. It is therefore possible to solve for x in the relation

$$\frac{\mathrm{d}f}{\mathrm{d}x} = p. \tag{B.19}$$

We would like to use the derivative p instead of x as an independent variable. If we solve for x in eq. (B.19) and the resulting expression of x in $f(x)$, we obtain indeed a function of p. However, we would obtain the same function if, instead of starting from $f(x)$, we started from a function $f_1(x) = f(x + C)$, where C is an arbitrary constant. Therefore this procedure discards important information on f. For example, if f is a function of x and other variables, C is an arbitrary function of all the variables but x. This ambiguity is removed by the Legendre transform, which allows us to define a function $g(p)$ that is equivalent to $f(x)$, in the sense that its knowledge allows us to recover precisely the original $f(x)$.

With our hypotheses, the function $f(x)$ is convex. We can represent it by a curve in the (x, f) plane. The same curve is also identified as the envelope of its tangents. Given a tangent of slope q at the point x, its intercept with the y axis is given by

$$\phi(p) = f(x) - p\,x. \tag{B.20}$$

Knowledge of $\phi(q)$ allows one to recover $f(x)$. Thus $\phi(p)$ is a good candidate for our task. However, it is more convenient to define the **Legendre transform** of f as the *opposite* of ϕ:

$$g(p) = p\,x - f(x), \tag{B.21}$$

where x is expressed as a function of q by solving eq. (B.19). This expression can also be obtained by the simple geometrical construction shown in fig. B.2.

With this convention, we have

$$\frac{\mathrm{d}g}{\mathrm{d}p} = x. \tag{B.22}$$

Moreover, the second derivative of g with respect to p is positive:

$$\frac{\mathrm{d}^2 g}{\mathrm{d}p^2} = \frac{\mathrm{d}x}{\mathrm{d}p} = \left(\frac{\mathrm{d}p}{\mathrm{d}x}\right)^{-1} > 0. \tag{B.23}$$

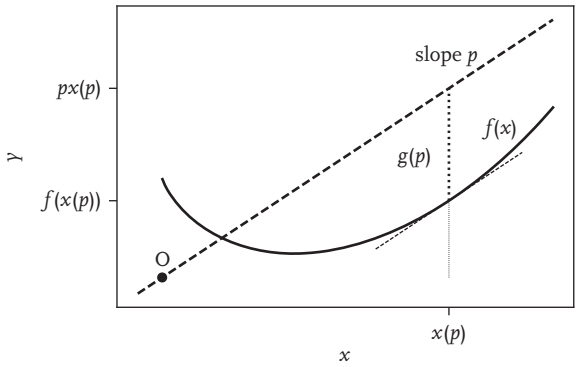


FIGURE B.2. Legendre transform of the function $f(x)$. To obtain the Legendre transform, we draw a line through the origin O with slope p. We then take the point $x(p)$ at which the vertical distance from the graph of f to this line (considered positive if $f(x)$ lies below the line, and negative if it lies above it) is maximized. Then $g(p)$ is given by the maximal value of this distance (with sign). If the tangent to the graph in $(x(p), f(x(p)))$ exists, it has slope p.

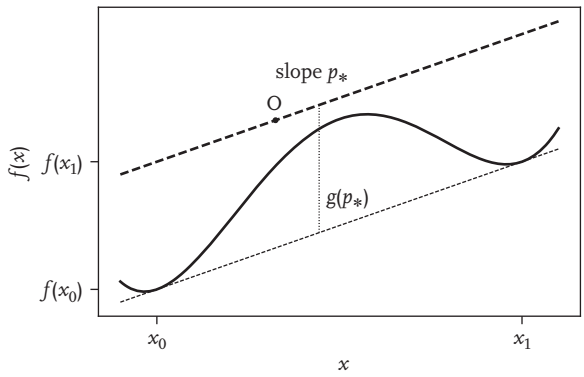


FIGURE B.3. Double tangent construction on a nonconvex, but differentiable, function $f(x)$. The slope of the double tangent is equal to p_*. To $p = p_*$, and thus to the value $g(p_*)$ of the Legendre transform, corresponds the whole interval $x_0 \leq x \leq x_1$.

We can thus apply the same transformation to $g(p)$. By eq. (B.21), this procedure yields $f(x)$ again. Therefore, under our assumptions, the Legendre transform is an involution.

If the function $f(x)$ is not convex (but still twice differentiable), the envelope of the tangents will be the convex envelope of the graph of $f(x)$, rather than the graph itself. This convex envelope is still given by the construction shown in fig. B.2. In particular, if $f(x)$ is not convex, some tangents have more than one contact point with the graph, as shown in fig. B.3. In these cases, eq. (B.19) has more than one solution, and, by eq. (B.22), there is a whole range of tangents to the $g(p)$ graphs for the corresponding values of p. Thus $g(p)$ exhibits angular points. This situation is shown in fig. B.4. When this happens, the Legendre transform is not an involution: applying it twice we recover the convex envelope of the graph of $f(x)$ rather than the graph itself. For example, for the function shown in fig. B.3, one obtains beween x_0 and x_1 the linear function exhibited with dashed lines rather than $f(x)$.

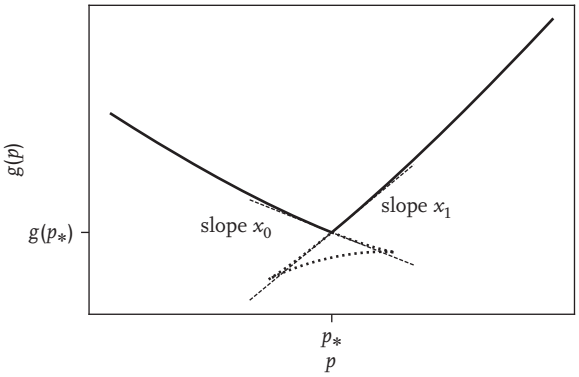


FIGURE B.4. Parametric plot of $x\,p(x) - f(x)$ vs. $p(x)$ for the function shown in fig. B.3, exhibiting a crossing for $p = p_*$. The Legendre transform of $f(x)$, given by the variational principle shown in eq. (B.25), is represented by the full line and exhibits an angular point at p_*, where the tangents have slopes x_0 and x_1, respectively. The dotted line corresponds to the values of x between x_0 and x_1, where the Legendre transform remains given by $g(p_*)$.

B.4 Properties of the Legendre Transformation

The Legendre transformation from the function $f(x)$ to its Legendre transform $g(q)$ has the following properties:

Linearity: If $f = \lambda_1 f_1 + \lambda_2 f_2$, then its Legendre transform is given by $g = \lambda_1 g_1 + \lambda_2 g_2$, where g_i is the Legendre transform of f_i for $i = 1, 2$.

Convexity: If f is convex with respect to x, g is convex with respect to p.

Variational principle: Since $f(x)$ is convex, its graph lies always above any tangent to it. Thus if x_0 is such that $f'(x_0) = p$, we have

$$g(p) = p\,x_0 - f(x_0) \geq p\,x - f(x), \qquad \forall x. \tag{B.24}$$

We thus obtain the following expression of $g(p)$:

$$g(p) = \max_x \big(p\,x - f(x)\big), \tag{B.25}$$

which expresses the construction shown in fig. B.2.

As a consequence of the variational principle, when the function $f(x)$ is not convex, the construction yields the same value of $g(p)$ for a whole interval of values of x, as shown in fig. B.3. Then the parametric plot of $(p(x), x\,p(x) - f(x))$ crosses itself, and only the larger values of $x\,p(x) - f(x))$ contribute to $g(p)$. This is shown in fig. B.4, where the Legendre transform $g(p)$ exhibits an angular point at $p = p_*$.

B.5 Lagrange Multipliers

Let us suppose that we want to look for the maxima and minima of a function $f(x, y)$ of the two variables x and y. Only, instead of looking for the *unconstrained* maxima and

minima—in other words, within a domain of the xy plane—we want to identify them among the points that satisfy a condition of the kind

$$\phi(x, y) = 0, \tag{B.26}$$

where $\phi(x, y)$ is a certain function. Denoting by $\mathcal{C}$ the locus of the points that satisfy this condition, let us look for the maxima and minima of f in $\mathcal{C}$. Problems of this kind are called **constrained** extremum problems, where eq. (B.26) expresses the constraint that the solution must satisfy.

Let us assume that we are able to solve eq. (B.26) explicitly with respect to y, obtaining $y(x)$ such that

$$\phi(x, y(x)) = 0. \tag{B.27}$$

Then our problem becomes an unconstrained problem, with x as the only unknown, of looking for the extrema of the function $\psi(x)$ defined by

$$\psi(x) = f(x, y(x)). \tag{B.28}$$

Let x_0 be one extremum of ψ, where we obviously have $\psi'(x_0) = 0$. We thus obtain

$$0 = \psi'(x_0) = \left.\frac{\partial f}{\partial x}\right|_{x_0, y(x_0)} + \left.\frac{\partial f}{\partial y}\right|_{x_0, y(x_0)} \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x_0}. \tag{B.29}$$

On the other hand, by taking the derivative of eq. (B.26), we obtain

$$\left.\frac{\partial \phi}{\partial x}\right|_{x_0, y(x_0)} + \left.\frac{\partial \phi}{\partial y}\right|_{x_0, y(x_0)} \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x_0} = 0. \tag{B.30}$$

Let us multiply this equation by an undetermined real number λ, and subtract both sides of eq. (B.30) from eq. (B.29). We obtain

$$0 = \left.\left(\frac{\partial f}{\partial x} - \lambda \frac{\partial \phi}{\partial x}\right)\right|_{x_0, y(x_0)} + \left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_{x_0} \left.\left(\frac{\partial f}{\partial y} - \lambda \frac{\partial \phi}{\partial y}\right)\right|_{x_0, y(x_0)}. \tag{B.31}$$

If we choose λ such that the second term of this expression vanishes, this implies that the first term also vanishes. But $\partial f/\partial x - \lambda \partial \phi/\partial x$ is the partial derivative of $\Phi(x, y; \lambda) = f(x, y) - \lambda \phi(x, y)$ with respect to x (and analogously for the case of y). In other words, there is a value of the constant λ such that the unconstrained extremum of $\Phi(x, y; \lambda)$ with respect to x and y is the constrained extremum of $f(x, y)$ on $\mathcal{C}$. The constant λ is called the **Lagrange multiplier**.

This defines a strategy for identifying the constrained extrema. One searches for the unconstrained extrema of $\Phi(x, y; \lambda)$ as a function of λ. As λ varies, these extrema trace a curve $\mathcal{C}' = \{x = x_0(\lambda), y = y_0(\lambda)\}$. One then looks at the points where $\mathcal{C}'$ crosses the curve $\mathcal{C}$ where the constraint $\phi(x, y) = 0$ is satisfied. The procedure is sketched in fig. B.5. It is easy to generalize the previous considerations to cases with several variables and to define the sufficient conditions for this method to work.

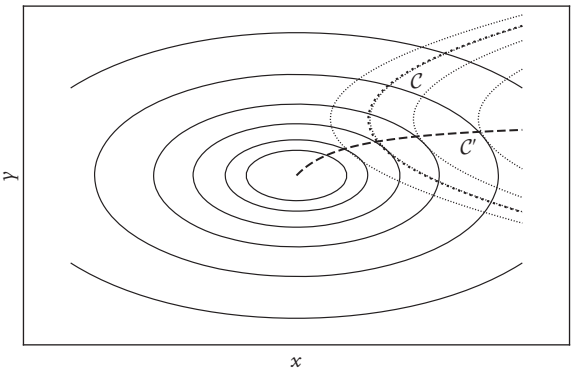


FIGURE B.5. Constrained optimization in the (x, y) plane. The full lines correspond to $f(x, y) = \text{const}$. The thick dotted line represents the line $\mathcal{C}$, where the constraint $\phi(x, y) = 0$ is satisfied, while the thin dotted lines correspond to $\phi(x, y) = \text{const}$. The dashed line $\mathcal{C}'$ is the locus of optima of $\Phi(x, y; \lambda) = f(x, y) - \lambda\phi(x, y)$ as λ varies. The constrained optimum lies at the crossing of $\mathcal{C}$ and $\mathcal{C}'$, where the level line of $f(x, y)$ is tangent to $\mathcal{C}$.

We can now see the connection between Lagrange multipliers and Legendre transforms. By the variational principle of the Legendre transform, $g(p, y)$ is the minimum over x of $\Phi(x, y; p) = f(x, y) - p\,x$. This corresponds to the constrained minimum of $f(x, y)$ on the line $x = \text{const.}$, where p is a Lagrange multiplier that fixes the value of x.

Appendix C Partial and Functional Derivatives

C.1 Partial Derivatives

Here we recall some useful properties of partial derivatives.

Consider a function $f(x, y)$ of two real variables x, y. We denote by $\partial f/\partial x)_y$ the partial derivative of f with respect to x, when y is kept constant. Formally, defining the function $F(x) = f(x, y_0)$, where y_0 is a constant, we have

$$\frac{\partial f(x, y_0)}{\partial x}\Big)_y = \frac{\mathrm{d}F(x)}{\mathrm{d}x}. \tag{C.1}$$

The definition is easily generalized to the case of more independent variables. We sometimes use the notations

$$\frac{\partial f}{\partial x}\Big)_y = \partial_x f(x, y); \tag{C.2}$$

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \partial_{y,x} f(x, y). \tag{C.3}$$

We also use the notation

$$\frac{\partial f}{\partial x}\bigg|_{x_0, y_0} = \partial_x f(x_0, y_0) \tag{C.4}$$

to denote the partial derivative $\partial f/\partial x)_y$, evaluated for $x = x_0$ and $y = y_0$.

The partial derivative satisfies the obvious relation

$$\frac{\partial f}{\partial x}\Big)_y = \left[\frac{\partial x}{\partial f}\Big)_y\right]^{-1}. \tag{C.5}$$

Assume, more generally, that the three real variables x, y, and z satisfy a relation of the form

$$f(x, y, z) = 0. \tag{C.6}$$

We then have

$$\mathrm{d}f = \frac{\partial f}{\partial x}\bigg)_{y,z} \mathrm{d}x + \frac{\partial f}{\partial y}\bigg)_{x,z} \mathrm{d}y + \frac{\partial f}{\partial z}\bigg)_{x,y} \mathrm{d}z = 0. \tag{C.7}$$

This implies, for example,

$$\frac{\partial z}{\partial x}\bigg)_{y} = -\frac{\partial f/\partial x)_{y,z}}{\partial f/\partial z)_{x,y}}. \tag{C.8}$$

Now, let $\phi(x, y)$ be some function of its arguments. We then have, for example,

$$\frac{\partial z}{\partial x}\bigg)_{\phi} = \frac{\partial z}{\partial x}\bigg)_{y} + \frac{\partial z}{\partial y}\bigg)_{x} \frac{\partial y}{\partial x}\bigg)_{\phi} = \frac{\partial z}{\partial x}\bigg)_{y} - \frac{\partial z}{\partial y}\bigg)_{x} \frac{\partial \phi/\partial x)_{y}}{\partial \phi/\partial y)_{x}}. \tag{C.9}$$

Suppose that $\phi(x, y)$ can be differentiated twice with continuity with respect to its arguments. Then the mixed derivatives are equal:

$$\left.\frac{\partial^2 \phi}{\partial x \partial y}\right|_{x,y} = \left.\frac{\partial^2 \phi}{\partial y \partial x}\right|_{x,y}, \qquad \forall x, y. \tag{C.10}$$

We have, indeed, for a given (x_0, y_0),

$$\begin{aligned}
\phi(x_0 + \mathrm{d}x, y_0 + \mathrm{d}y) &\simeq \phi(x_0, y_0) + \partial_x \phi(x_0, y_0)\, \mathrm{d}x + \partial_y \phi(x_0 + \mathrm{d}x, y_0)\, \mathrm{d}y \\
&\simeq \phi(x_0, y_0) + \partial_x \phi(x_0, y_0)\, \mathrm{d}x + \partial_y \phi(x_0, y_0)\, \mathrm{d}y + \partial_{x,y} \phi(x_0, y_0)\, \mathrm{d}x\, \mathrm{d}y \\
&\simeq \phi(x_0, y_0) + \partial_y \phi(x_0, y_0)\, \mathrm{d}y + \partial_y \phi(x_0, y_0 + \mathrm{d}y)\, \mathrm{d}x \\
&\simeq \phi(x_0, y_0) + \partial_y \phi(x_0, y_0)\, \mathrm{d}y + \partial_x \phi(x_0, y_0)\, \mathrm{d}x + \partial_{y,x} \phi(x_0, y_0)\, \mathrm{d}y\, \mathrm{d}x.
\end{aligned} \tag{C.11}$$

Therefore

$$\partial_{x,y} \phi(x_0, y_0) = \partial_{y,x} \phi(x_0, y_0), \qquad \forall x_0, y_0. \tag{C.12}$$

An expression of the form

$$đX = \phi(x, y)\, \mathrm{d}x + \chi(x, y)\, \mathrm{d}y \tag{C.13}$$

is called a **differential form**. In general, it is not the differential of a function of x and y. A necessary condition for this follows from the equality of mixed derivatives:

$$\partial_y \phi(x, y) = \partial_x \chi(x, y), \qquad \forall x, y. \tag{C.14}$$

If the form satisfies this condition, it is denoted as

$$\mathrm{d}X = \phi(x, y)\, \mathrm{d}x + \chi(x, y)\, \mathrm{d}y, \tag{C.15}$$

and one says that $\mathrm{d}X$ is an **exact differential**. (However, it is not necessarily the differential of a one-valued function of x and y.) It may be the case that there is a function $\lambda(x, y)$ such that $\lambda\, đX$ is an exact differential. If this is the case, the function $\lambda(x, y)$ is called an **integrating factor**. As an example, consider the differential form

$$đX = \phi\, \mathrm{d}x + \chi\, \mathrm{d}y, \tag{C.16}$$

where

$$\phi = -y, \qquad \chi = x. \tag{C.17}$$

Since $\partial_y \phi = -1 \neq \partial_x \chi = 1$, this is not an exact differential. Introducing the integrating factor

$$\lambda(x, y) = \frac{1}{x^2 + y^2}, \tag{C.18}$$

we obtain the exact differential

$$\mathrm{d}Y = -\frac{y}{x^2 + y^2}\, \mathrm{d}x + \frac{x}{x^2 + y^2}\, \mathrm{d}y. \tag{C.19}$$

If the independent variable is a vector, such as a position vector $\boldsymbol{r} = (r_x, r_y, r_z)$, we denote by $\partial f / \partial \boldsymbol{r}$ the vector formed by the partial derivatives of f with respect to the components of $\boldsymbol{r}$:

$$\frac{\partial f}{\partial \boldsymbol{r}} = \left(\frac{\partial f}{\partial r_x}, \frac{\partial f}{\partial r_y}, \frac{\partial f}{\partial r_z} \right). \tag{C.20}$$

C.2 Functional Derivatives

The notion of partial derivatives generalizes quite naturally to **functional derivatives**. In this case one considers a **functional**, that is, a quantity $F[f]$ that depends on the whole function $f(x)$, where x belongs to some domain, and $f(x)$ satisfies some regularity and boundary conditions. A trivial example is given by

$$F[f] = \int_a^b \mathrm{d}x f(x), \tag{C.21}$$

where $f(x)$ is defined in the interval $\mathcal{I} = a \leq x \leq b$ and $f(x)$ is a continuous function in the interval. By convention, the argument of a functional is enclosed in brackets. A slightly more general case is given by

$$F_p[f] = \int_a^b \mathrm{d}x\, p\,(x, f(x)), \tag{C.22}$$

where $p(x, y)$ is a sufficiently regular function defined in $\mathcal{I} \times \mathbb{R}$.

Suppose that we wish to evaluate the change in $F_p[f]$ when $f(x)$ changes by a small quantity $\delta f(x)$ for all $x \in \mathcal{I}$. We can discretize the problem by sampling the values of $f(x) + \delta f(x)$

at the $\mathcal{N}+1$ points $x_k = a + k[(b-a)/\mathcal{N}]$, $k = 0, 1, \ldots, N$, and by evaluating a discretized approximation $\mathcal{F}(\phi_0, \ldots, \phi_\mathcal{N})$ to $F_p[f+\delta f]$, where $\phi_k = f(x_k)$. We obtain

$$\mathcal{F} = \sum_{k=0}^{N} p(x_k, \phi_k)\, \Delta x_k, \tag{C.23}$$

where $\Delta x_k = (b-a)/\mathcal{N}$. This is a function of the $\mathcal{N}+1$ variables $(\phi_0, \ldots, \phi_\mathcal{N})$. We then have, for small enough $|\delta f(x)|$,

$$\mathcal{F}[f+\delta f] - \mathcal{F}[f] \simeq \sum_{k=0}^{\mathcal{N}} \left.\frac{\partial p(x,y)}{\partial y}\right|_{y=\phi_k} \delta f(x_k)\, \Delta x_k, \tag{C.24}$$

up to terms of higher order in the $\delta f(x_k)$. If $\delta f(x)$ is continuous, this expression converges, as $\mathcal{N} \to \infty$, to

$$F[f+\delta f] - F[f] \simeq \int_a^b \mathrm{d}x \left.\frac{\partial p(x,y)}{\partial y}\right|_{y=f(x)} \delta f(x). \tag{C.25}$$

This relation is usually written in the form

$$\frac{\delta F}{\delta f(x)} = \frac{\partial p}{\partial f(x)}. \tag{C.26}$$

The expression on the left-hand side is the **functional derivative** of the functional $F[f]$ with respect to its argument. Our result can thus be written in the form

$$F[f+\delta f] - F[f] \simeq \int_a^b \mathrm{d}x\, \frac{\delta F}{\delta f(x)}\, \delta f(x). \tag{C.27}$$

Let us now look at the more general (and important) case where the functional is obtained as an integral of a function that depends, beyond x and $f(x)$, on the derivative $f'(x)$ of $f(x)$, evaluated at the running point x. In this case we also require that the derivative $f'(x)$ is continuous. We then have

$$F_p[f] = \int_a^b \mathrm{d}x\, p(x, f(x), f'(x)), \tag{C.28}$$

where $p(x,y,z)$ is a sufficiently regular function defined in $\mathcal{I} \times \mathbb{R}^2$. Using the same approach as before, we obtain

$$F[f+\delta f] - F[f] \simeq \int_a^b \mathrm{d}x \left[\frac{\partial p}{\partial f(x)}\delta f(x) + \frac{\partial p}{\partial f'(x)}\, \delta f'(x)\right]. \tag{C.29}$$

We now have, integrating by parts, and assuming that the second derivative $f''(x)$ of $f(x)$ exists and is continuous,

$$\int_a^b \mathrm{d}x\, \frac{\partial p}{\partial f'(x)}\, \delta f'(x) = \left[\frac{\partial p}{\partial f'(x)}\, \delta f(x)\right]_a^b - \int_a^b \mathrm{d}x\, \delta f(x)\, \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial p}{\partial f'(x)}, \tag{C.30}$$

where

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial p}{\partial f'(x)} = \left[\frac{\partial p}{\partial x} + f'(x)\frac{\partial p}{\partial y} + f''(x)\frac{\partial p}{\partial z}\right]\Bigg|_{x,y=f(x),z=f'(x)}. \tag{C.31}$$

If we further assume that $\delta f(a) = \delta f(b) = 0$, we obtain the expression of the functional derivative in the form

$$\frac{\delta F}{\delta f(x)} = \frac{\partial p}{\partial f(x)} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial p}{\partial f'(x)}. \tag{C.32}$$

The functional derivative satisfies the usual properties of the derivative.

Linearity:

$$\frac{\delta(\lambda F + \mu G)}{\delta f(x)} = \lambda\frac{\delta F}{\delta f(x)} + \mu\frac{\delta G}{\delta f(x)}; \tag{C.33}$$

Product rule:

$$\frac{\delta(FG)}{\delta f(x)} = \frac{\delta F}{\delta f(x)}G + F\frac{\delta G}{\delta f(x)}; \tag{C.34}$$

Chain rule:

$$\frac{\delta F[g(f)]}{\delta f(x)} = \frac{\delta F}{\delta g(f(x))}\frac{\mathrm{d}g(f(x))}{\mathrm{d}f(x)}; \qquad \frac{\delta g(F[f])}{\delta f(x)} = \frac{\mathrm{d}g(F[f])}{\mathrm{d}F[f]}\frac{\delta F[f]}{\delta f(x)}. \tag{C.35}$$

These considerations generalize easily to the case where the independent variable $\boldsymbol{x}$ varies in a domain of dimension $n > 1$. In particular, for the functional

$$F[f] = \int \mathrm{d}\boldsymbol{x}\, p(\boldsymbol{x}, f(\boldsymbol{x}), \nabla f(\boldsymbol{x})), \tag{C.36}$$

the functional derivative is given by

$$\frac{\delta F}{\delta f(\boldsymbol{x})} = \frac{\partial p}{\partial f(\boldsymbol{x})} - \nabla\cdot\frac{\partial p}{\partial \nabla f(\boldsymbol{x})}, \tag{C.37}$$

provided the variation $\delta f(\boldsymbol{x})$ vanishes on the boundary of the domain spanned by $\boldsymbol{x}$.

Appendix D Fourier Transforms

D.1 Fourier Series

Let $f(x)$ be a periodic function of x, with period L. This means that

$$f(x+L)=f(x), \qquad \forall x. \tag{D.1}$$

Notice that, for any integer n, the functions $\cos(2\pi nx/L)$ and $\sin(2\pi nx/L)$ satisfy this condition, and therefore so does any linear combination of them of the form

$$f(x;a,b)=\sum_n \left[a_n \cos\left(\frac{2\pi nx}{L}\right)+b_n \sin\left(\frac{2\pi nx}{L}\right)\right], \tag{D.2}$$

where $a=(a_n)=(a_0,a_1,a_2,\ldots)$ and $b=(b_n)=(0,b_1,b_2,\ldots)$. It is natural to ask which sequences a and b are such that $f(x;a,b)=f(x)$, $\forall x$, for a given $f(x)$.

Let us assume that these sequences exist, and observe that, if $n,m\geq 0$,

$$\begin{aligned}
&\int_{-L/2}^{L/2} \mathrm{d}x\, \cos\left(\frac{2\pi nx}{L}\right)\cos\left(\frac{2\pi mx}{L}\right)=\begin{cases} L, & \text{if } n=m=0;\\ L/2, & \text{if } n=m>0;\\ 0, & \text{if } n\neq m.\end{cases}\\
&\int_{-L/2}^{L/2} \mathrm{d}x\, \sin\left(\frac{2\pi nx}{L}\right)\sin\left(\frac{2\pi mx}{L}\right)=\begin{cases} 0, & \text{if } n=0 \text{ or } m=0;\\ L/2, & \text{if } n=m>0;\\ 0, & \text{if } n\neq m.\end{cases}\\
&\int_{-L/2}^{L/2} \mathrm{d}x\, \sin\left(\frac{2\pi nx}{L}\right)\cos\left(\frac{2\pi mx}{L}\right)=0, \qquad \forall n,m.
\end{aligned} \tag{D.3}$$

Thus, if $f(x)=f(x;a,b)$, where $f(x;a,b)$ is given by eq. (D.2), and we can interchange the integration with the sum, we have

$$\int_{-L/2}^{L/2} \mathrm{d}x f(x) \cos\left(\frac{2\pi nx}{L}\right) = \begin{cases} a_0 L, & \text{if } n = 0; \\ a_n L/2, & \text{if } n > 0; \end{cases}$$
$$\int_0^L \mathrm{d}x f(x) \sin\left(\frac{2\pi nx}{L}\right) = b_n L/2, \qquad n > 0. \tag{D.4}$$

This allows us to obtain the **cosine and sine coefficients** (a_n, b_n):

$$\begin{aligned} a_0 &= \frac{1}{L}\int_{-L/2}^{L/2} \mathrm{d}x f(x); \\ a_n &= \frac{2}{L}\int_{-L/2}^{L/2} \mathrm{d}x f(x) \cos\left(\frac{2\pi nx}{L}\right), \qquad n > 0; \\ b_n &= \frac{2}{L}\int_{-L/2}^{L/2} \mathrm{d}x f(x) \sin\left(\frac{2\pi nx}{L}\right), \qquad n > 0. \end{aligned} \tag{D.5}$$

Notice that, if $f(-x) = f(x)$, $\forall x$, we have $b_n = 0$, $\forall n$, whereas if $f(-x) = -f(x)$, $\forall x$, we have $a_n = 0$, $\forall n$. In the first case, we have a **cosine series** representation, and in the second case, a **sine series** one.

We have thus obtained a mapping from a continuous function $f(x)$, defined on the whole real axis, to two sequences $a = (a_n)$ and $b = (b_n)$, defined for $n \geq 0$. A more compact representation can be obtained by exploiting Euler's identity,

$$\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i} \sin x. \tag{D.6}$$

We can thus write

$$f(x, n) = \sum_{n=-\infty}^{+\infty} c_n \, \mathrm{e}^{2\pi \mathrm{i} nx/L}. \tag{D.7}$$

By exploiting the relation

$$\int_{-L/2}^{L/2} \mathrm{d}x \, \mathrm{e}^{-2\pi \mathrm{i}(n-m)x/L} = \begin{cases} L, & \text{if } n = m; \\ 0, & \text{otherwise} \end{cases} \tag{D.8}$$

we obtain the expression for the **Fourier coefficients** c_n:

$$c_n = \frac{1}{L}\int_{-L/2}^{L/2} \mathrm{d}x f(x) \, \mathrm{e}^{-2\pi \mathrm{i} nx/L}. \tag{D.9}$$

Since $f(x)$ is real, we obtain the following symmetry relation:

$$c_{-n} = c_n^*, \qquad \forall n. \tag{D.10}$$

One can check that

$$c_n = \frac{a_n - \mathrm{i} b_n}{2}, \qquad n \geq 0, \tag{D.11}$$

and use eq. (D.10) to express c_n for $n < 0$.

The representation of the periodic function $f(x)$ in terms of the series (D.2) or eq. (D.7) is called the **Fourier series** representation of $f(x)$.

It is useful to introduce the **wavenumber** k_n, defined by

$$k_n = \frac{2\pi n}{L}. \tag{D.12}$$

We then have

$$f(x) = \sum_n c_n \, \mathrm{e}^{\mathrm{i}k_n x}, \tag{D.13}$$

where

$$c_n = \frac{1}{L} \int_{-L/2}^{L/2} \mathrm{d}x f(x) \, \mathrm{e}^{-\mathrm{i}k_n x}. \tag{D.14}$$

One can apply a Fourier series to functions $f(x)$ that identically vanish outside an interval. Indeed, if we evaluate the Fourier coefficients via (D.9) on any interval where $f(x)$ is not identically zero, we obtain a series that reproduces $f(x)$ within the interval, and extends the function periodically outside it.

D.2 Fourier Transform

Let $f(x)$ be a real-valued function such that

$$\int_{-\infty}^{+\infty} \mathrm{d}x \, \big|f(x)\big| < \infty. \tag{D.15}$$

One then says that $f(x)$ is **integrable** on the real axis. Choosing an arbitrarily large value of L, we can then define the analog of c_n, defined in (D.14), by

$$f_L(k_n) = Lc_n = \int_{-L/2}^{+L/2} \mathrm{d}x f(x) \, \mathrm{e}^{-\mathrm{i}k_n x}. \tag{D.16}$$

In our hypotheses, the integral

$$f(k) = \lim_{L\to\infty} \int_{-L/2}^{+L/2} \mathrm{d}x f(x) \, \mathrm{e}^{-\mathrm{i}kx} = \lim_{L\to\infty} f_L(k) \tag{D.17}$$

converges for all real values of the wavenumber k. Then the function $f(k)$ is called the **Fourier transform** of $f(x)$. If $f(k)$ is itself integrable on the real axis, one has

$$\begin{aligned} f(x) &= \lim_{L\to\infty} \frac{1}{L} \sum_n f_L(k_n) \, \mathrm{e}^{\mathrm{i}k_n x} = \lim_{L\to\infty} \frac{1}{L} \cdot \frac{L}{2\pi} \int \mathrm{d}k f_L(k) \, \mathrm{e}^{\mathrm{i}kx} \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k f(k) \, \mathrm{e}^{\mathrm{i}kx}. \end{aligned} \tag{D.18}$$

One expresses this relation by saying that $f(x)$ is the **inverse Fourier transform** of $f(k)$. The **Fourier transformation** $\mathcal{F} : f(x) \longrightarrow f(k)$ is linear, and has a number of important properties. In particular, the Fourier transform of the derivative $f'(x)$ is given by $\mathrm{i}kf(k)$. Since the

transformation and its inverse have similar expressions, this also means that the Fourier transform of $xf(x)$ is given by $-if'(k)$.

Given two functions $f(x)$ and $g(x)$, define their **convolution** $f * g$ by

$$(f * g)(x) = \int_{-\infty}^{+\infty} \mathrm{d}y f(x-y)\, g(y). \tag{D.19}$$

We then obtain

$$\begin{aligned}
\mathcal{F}(f * g) &= \int_{-\infty}^{+\infty} \mathrm{d}x\, (f * g)(x)\, \mathrm{e}^{-\mathrm{i}kx} \\
&= \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y f(x-y)\, g(y)\, \mathrm{e}^{-\mathrm{i}kx} \\
&= \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y f(x-y)\, \mathrm{e}^{-\mathrm{i}k(x-y)}\, g(y)\, \mathrm{e}^{-\mathrm{i}ky} \\
&= \int_{-\infty}^{+\infty} \mathrm{d}z f(z)\, \mathrm{e}^{-\mathrm{i}kz} \int_{-\infty}^{+\infty} \mathrm{d}y\, g(y)\, \mathrm{e}^{-\mathrm{i}ky} \\
&= f(k) \cdot g(k).
\end{aligned} \tag{D.20}$$

Thus convolution maps into the product of the Fourier transforms, and, vice versa, the product of two functions maps into the convolution of their transforms.

Consider now the integral defined by

$$\mathcal{I} = \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y f(x) K(x-y) g(y), \tag{D.21}$$

where $K(x)$ is defined for all $x \in \mathbb{R}$. We can express $K(x-y)$ in terms of its Fourier transform $K(k)$:

$$K(x-y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}k(x-y)}\, K(k). \tag{D.22}$$

Substituting in the expression of $\mathcal{I}$, we obtain

$$\begin{aligned}
\mathcal{I} &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y\, \mathrm{e}^{\mathrm{i}kx} f(x)\, K(k)\, \mathrm{e}^{-\mathrm{i}ky} g(y) \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k f(-k)\, K(k)\, g(k).
\end{aligned} \tag{D.23}$$

The Fourier transform of a Gaussian is itself a Gaussian. We have, indeed, when

$$f(x) = \exp\left[-\frac{(x-x_0)^2}{2\sigma^2}\right], \tag{D.24}$$

that the Fourier transform is given by

$$\begin{aligned} f(k) &= \int_{-\infty}^{+\infty} \mathrm{d}x \, \exp\left[-\frac{(x-x_0)^2}{2\sigma^2}\right] \mathrm{e}^{-\mathrm{i}kx} \\ &= \int_{-\infty}^{+\infty} \mathrm{d}x \, \exp\left[-\frac{x^2 - 2xx_0 + x_0^2 + 2\mathrm{i}k\sigma^2 x}{2\sigma^2}\right] \\ &= \int_{-\infty}^{+\infty} \mathrm{d}x \, \exp\left\{-\frac{[x-(x_0-\mathrm{i}\sigma^2 k)]^2 + 2\mathrm{i}\sigma^2 k x_0 + \sigma^4 k^2}{2\sigma^2}\right\} \\ &= \sqrt{2\pi\sigma^2}\, \mathrm{e}^{-\sigma^2 k^2/2 - \mathrm{i}kx_0}. \end{aligned} \tag{D.25}$$

The width of the Fourier transform is equal to the inverse of the width of the original Gaussian. Similar uncertainty relations hold for more general forms of $f(x)$ and their transforms.

Fourier series and Fourier transforms are naturally extended to functions defined in $\mathbb{R}^d$. Given the function $f(\boldsymbol{r})$ of the position vector $\boldsymbol{r} = (r_1, \ldots, r_d)$, its Fourier transform is defined by

$$f(\boldsymbol{k}) = \int \mathrm{d}\boldsymbol{r} f(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{D.26}$$

where the integral runs over the whole $\mathbb{R}^d$ space spanned by $\boldsymbol{r}$. The Fourier antitransform is then given by

$$f(\boldsymbol{r}) = \frac{1}{(2\pi)^d} \int \mathrm{d}\boldsymbol{k} f(\boldsymbol{k})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{D.27}$$

where the integral spans the whole $\boldsymbol{k}$ space.

D.3 Functions Defined on a Lattice: Brillouin Zones

When the function $f(x)$ is defined on a lattice, its Fourier transform exhibits a periodic structure. Let $f(x)$ be defined on the points $x = -L, -L+1, \ldots, L-1, L$, where L is an integer. The Fourier transform is defined by

$$f(k) = \sum_{x=-L}^{+L} f(x)\, \mathrm{e}^{-\mathrm{i}kx}. \tag{D.28}$$

In order to extend the domain of f to all integers, the wavenumber k must be of the form

$$k_n = \frac{2\pi n}{2L+1}, \qquad n \in \mathbb{Z}. \tag{D.29}$$

We see, however from eq. (D.28) that, since x has integer values, one has

$$f(k+q) = f(k) \tag{D.30}$$

whenever

$$q = 2\pi m, \qquad m \in \mathbb{Z}. \tag{D.31}$$

We thus obtain the same value of the Fourier transform for wavenumbers k that form an infinite lattice. It is then convenient to constrain the values of k in order to eliminate this redundancy. In our case, this can be done by looking at the $(2L+1)$ values of k of smallest magnitude:

$$k \in \left\{ -2\pi \frac{L}{2L+1}, -2\pi \frac{L-1}{2L+1}, \ldots, 2\pi \frac{L-1}{2L+1}, 2\pi \frac{L}{2L+1} \right\}. \tag{D.32}$$

As L increases, the number of acceptable wavenumbers in the interval $[-\pi, \pi]$ also increases, so that in the $L \to \infty$ limit, we can consider $f(k)$ to be defined on the whole real axis, repeating periodically with period 2π. We then have

$$f(x) = \frac{1}{2L+1} \sum_{n=-L}^{+L} f(k_n)\, \mathrm{e}^{\mathrm{i}k_n x}. \tag{D.33}$$

In the limit $L \to \infty$, since $\mathrm{d}k_n = k_{n+1} - k_n = 2\pi/(2L+1)$, we obtain an integral over the interval $-\pi \le k \le \pi$:

$$f(x) = \int_{-\pi}^{+\pi} \frac{\mathrm{d}k}{2\pi} f(k)\, \mathrm{e}^{\mathrm{i}kx}. \tag{D.34}$$

The argument generalizes to d-dimensional lattices. Let $f(\boldsymbol{r})$ be defined for vectors $\boldsymbol{r}$ of the form

$$\boldsymbol{r}(n) = \sum_{\alpha=1}^{d} n_\alpha \boldsymbol{b}_\alpha, \tag{D.35}$$

where

$$n = (n_1, \ldots, n_d), \qquad n_\alpha \in \mathbb{Z}, \quad \forall \alpha, \tag{D.36}$$

and the linearly independent vectors $\boldsymbol{b}_\alpha$ are the **basis vectors** of the lattice. The Fourier transform of $f(\boldsymbol{r})$ is defined by

$$f(\boldsymbol{k}) = \sum_n f\left(\boldsymbol{r}(n)\right)\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}(n)}. \tag{D.37}$$

The transform has the periodicity

$$f(\boldsymbol{k}+\boldsymbol{q}) = f(\boldsymbol{k}), \qquad \forall \boldsymbol{k}, \tag{D.38}$$

whenever $\boldsymbol{q}$ satisfies

$$\boldsymbol{b}_\alpha \cdot \boldsymbol{q} = 2\pi\, m_\alpha, \quad m_\alpha \in \mathbb{Z}, \qquad \forall \alpha. \tag{D.39}$$

The vectors $\boldsymbol{q}$ that satisfy these conditions form the **reciprocal lattice** to the original one. As a consequence, we can split the space spanned by the wave vector $\boldsymbol{k}$ into zones, such that knowledge of $f(\boldsymbol{k})$ in one zone allows one to obtain the value of $f(\boldsymbol{k})$ for arbitrary $\boldsymbol{k}$ by taking advantage of the fact that $f(\boldsymbol{k}) = f(\boldsymbol{k}+\boldsymbol{q})$ whenever $\boldsymbol{q}$ belongs to the reciprocal lattice. The zones are called **Brillouin zones**. The first Brillouin zone is the region of points in $\boldsymbol{k}$ space

that are closer to the origin of the reciprocal lattice than they are to any other reciprocal lattice points.

We then obtain the following expression of the Fourier antitransform:

$$f(\boldsymbol{r}) = \frac{\mathcal{V}}{(2\pi)^d} \int_{\mathrm{BZ}} \mathrm{d}\boldsymbol{k} f(\boldsymbol{k})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{D.40}$$

where the integral runs over the first Brillouin zone, and $\mathcal{V}$ is the volume of a primitive cell of the lattice on which $f(\boldsymbol{r})$ is defined.

D.4 Integral Representation of the Delta Function

The relation between a function and its Fourier transform can be summarized by the expression

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k\, \exp(-\mathrm{i}kx) = \delta(x), \tag{D.41}$$

where $\delta(x)$ is Dirac's delta "function." One has, indeed,

$$f(x_0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k f(k)\, \exp(\mathrm{i}kx_0), \tag{D.42}$$

where $f(k)$ is the Fourier transform of $f(x)$:

$$f(k) = \int_{-\infty}^{+\infty} \mathrm{d}x\, \exp(-\mathrm{i}kx) f(x). \tag{D.43}$$

One has therefore

$$\begin{aligned} f(x_0) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{d}k f(k)\, \exp(\mathrm{i}kx_0) \\ &= \int_{-\infty}^{+\infty} \mathrm{d}x f(x) \int_{-\infty}^{+\infty} \frac{\mathrm{d}k}{2\pi}\, \exp\left[-\mathrm{i}k(x - x_0)\right]. \end{aligned} \tag{D.44}$$

Since

$$f(x_0) = \int_{-\infty}^{+\infty} \mathrm{d}x f(x)\, \delta(x - x_0), \tag{D.45}$$

and the function $f(x)$ is arbitrary, the result follows.

Appendix E Saddle-Point Integration

E.1 Euler Integrals and the Saddle-Point Method

It is often necessary to calculate integrals of the form

$$I(N) = \int_a^b \mathrm{d}x \, \exp\left[-Nf(x)\right], \tag{E.1}$$

where the variable N assumes large positive values, and where the function $f(x)$ exhibits a minimum in the integration path. These integrals are known as Euler integrals and can be estimated by the **saddle point method**. Let us assume that $f(x)$ exhibits a minimum in x_0, with $a < x_0 < b$. Then, let us expand $f(x)$ into a Taylor series around x_0:

$$f(x) = f(x_0) + \frac{1}{2} f''(x_0)\,(x - x_0)^2 + \cdots . \tag{E.2}$$

The first-order term is absent since x_0 corresponds to a minimum, and moreover we have $f''(x_0) > 0$. By substituting this expression in the integral, and by setting $\Delta x = x - x_0$, we obtain

$$I(N) = \exp\left[-Nf(x_0)\right] \int_{a-x_0}^{b-x_0} \mathrm{d}(\Delta x) \, \exp\left[-\frac{1}{2} Nf''(x_0)\,\Delta x^2 + \cdots\right]. \tag{E.3}$$

The error one introduces by taking the lower integration limit to $-\infty$ and the higher one to $+\infty$ is exponentially small in N. Neglecting higher-order terms in the Taylor expansion, we obtain

$$\begin{aligned} I(N) &\approx \exp\left[-Nf(x_0)\right] \int_{-\infty}^{+\infty} \mathrm{d}(\Delta x) \, \exp\left[-\frac{1}{2} Nf''(x_0)\Delta x^2\right] \\ &= \exp\left[-Nf(x_0)\right] \sqrt{\frac{2\pi}{Nf''(x_0)}}. \end{aligned} \tag{E.4}$$

The notation "$\mathrm{lhs}(N) \approx \mathrm{rhs}(N)$" means that

$$\lim_{N\to\infty} \frac{\mathrm{lhs}(N)}{\mathrm{rhs}(N)} = 1, \tag{E.5}$$

where "lhs" and "rhs" respectively denote the left-hand side and the right-hand side of the equation. Taking into account the following terms in the Taylor expansion of $f(x)$, and explicitly integrating them, one obtains the formal power series

$$I(N) \approx \exp\left[-Nf(x_0)\right] \sqrt{\frac{2\pi}{Nf''(x_0)}} \left(1 + \sum_{k=1}^{\infty} \frac{I_k}{N^k}\right). \tag{E.6}$$

The series is not convergent, but is rather an **asymptotic series**, in the sense that

$$I(N) \Bigg/ \left(\exp\left[-Nf(x_0)\right] \sqrt{\frac{2\pi}{Nf''(x_0)}}\right) = 1 + \sum_{k=1}^{r} \frac{I_k}{N^k} + \mathrm{o}(N^{-r}), \tag{E.7}$$

where we recall that $\mathrm{o}(x)$ means that $\lim_{x\to 0} (\mathrm{o}(x)/x) = 0$. Although they are not convergent, asymptotic series often represent an excellent approximation, and may even be preferred to convergent series from a numerical point of view.

The method we have just outlined can be generalized to integrals in the complex plane. For simplicity let us continue to assume that N is positive, but we now assume that the function $f(z)$ has an extremal (a point where its derivative $f'(z)$ vanishes) in a point z_0 in the complex plane. If $f(z)$ is analytic, we can deform the integration path in such a manner as to make it pass through z_0. Setting $z = z_0 + \Delta z$ and $\Delta f(z) = f(z) - f(z_0)$, we have, for small values of $|\Delta z|$,

$$\begin{aligned} \Delta f(z) \simeq \frac{1}{2} f''(z_0)\, \Delta z^2 = \mathrm{Re}\, f''(z_0) \left(x^2 - y^2\right) - 2\, \mathrm{Im}\, f''(z_0)\, xy \\ + \mathrm{i}\left[\mathrm{Im}\, f''(z_0) \left(x^2 - y^2\right) + 2\, \mathrm{Re}\, f''(z_0)\, xy\right]. \end{aligned} \tag{E.8}$$

where we have set $\Delta z = x + \mathrm{i}y$. Apart from the constant factor $\exp\left[-Nf(z_0)\right]$, the absolute value of the integrand is given by $\exp\left[-N\,\mathrm{Re}\, f(z)\right]$. Let us represent the function $u = \mathrm{Re}\, f(z)$ as a surface in the space (x, y, z). Since the function u is harmonic, it cannot exhibit maximum or minimum points, therefore the point z_0 must be a saddle point. The most convenient path for the evaluation of the integral is the one for which the change in the value of $\mathrm{Re}\, f(z)$ is fastest. This can be achieved by rotating the path around z_0, that is, by multiplying Δz by a phase factor $\mathrm{e}^{\mathrm{i}\theta}$. In this way the square magnitude $\left|\Delta f\right|^2 = (\mathrm{Re}\, \Delta f)^2 + (\mathrm{Im}\, \Delta f)^2$ does not change. Therefore the path with the largest value of $\mathrm{Re}\, \Delta f$ corresponds to the one in which $\mathrm{Im}\, f(z)$ remains constant. Along this path, the integral reduces, up to a constant factor, to the form shown in eq. (E.1). In the (x, y, z) space, this path looks like the steepest way up and down a saddle point situated at $(0, 0, f(z_0))$, which justifies the name given to the method. Since the same path can be characterized by the condition $\mathrm{Im}\, f(z) = \mathrm{const.}$, which implies that the phase of the integrand $\exp\left[-Nf(z)\right]$ remains constant, the method is also known as the stationary-phase method.

E.2 The Euler Gamma Function

Let us consider the integral $\Gamma_E(z)$, defined (for positive z) by

$$\Gamma_E(z) = \int_0^\infty \mathrm{d}t\, t^{z-1}\, \mathrm{e}^{-t}. \tag{E.9}$$

We obviously have $\Gamma_E(1) = 1$, while an integration by parts shows that

$$z\Gamma_E(z) = \Gamma_E(z+1). \tag{E.10}$$

These relations imply that, if n is a positive integer,

$$\Gamma_E(n) = (n-1)!. \tag{E.11}$$

Therefore $\Gamma_E(z)$ is a generalization of the factorial. The integral converges in the whole right half of the complex plane, and therefore $\Gamma_E(x)$ can be extended, by means of the relation (E.10), to the entire complex plane, except for the points $z = 0, -1, -2, \ldots$, where it exhibits simple poles. We have thus defined **Euler's gamma function**.

For $z = 1/2$, the transformation $t = x^2$ connects our integral to the Gaussian integral, and one therefore obtains $\Gamma_E(1/2) = \sqrt{\pi}$. By applying the functional relation (E.10), we obtain the relation

$$\begin{aligned} \Gamma_E\left(n + \frac{1}{2}\right) &= \frac{(2n-1)(2n-3)\cdots}{2^n}\sqrt{\pi} \\ &= \frac{(2n-1)!!}{2^n}\sqrt{\pi}. \end{aligned} \tag{E.12}$$

Here $n!!$ is the semifactorial, which is equal to $2 \cdot 4 \cdot \ldots \cdot n$ if n is even and to $3 \cdot 5 \cdot \ldots \cdot n$ if n is odd.

We can now evaluate $\Gamma_E(z)$ for large values of z by the saddle-point method. We bring the power t^{z-1} to the exponent, by $t^{z-1} = \exp[(z-1)\ln t]$. Then the maximum of the exponent lies at $t = t_* = z - 1$. We thus have

$$\begin{aligned} \Gamma_E(z) &= \int_0^\infty \mathrm{d}t\, \exp[(z-1)\ln t - t] \\ &\approx \exp[(z-1)\ln t_* - t_*] \int_{-\infty}^{+\infty} \mathrm{d}(\Delta t)\, \exp\left(-\frac{z-1}{2t_*^2}\Delta t^2\right) \\ &= \exp[(z-1)\ln(z-1) - (z-1)]\, \sqrt{2\pi\,(z-1)}. \end{aligned} \tag{E.13}$$

Setting $z = n + 1$, we obtain the well-known **Stirling formula**:

$$n! = \Gamma_E(n+1) \approx \exp(n\ln n - n)\sqrt{2\pi n}. \tag{E.14}$$

This result can be extended to complex values of z with a large positive real part by deforming the path in the complex plane, as described in sec. E.1.

It is also useful to evaluate the following integral (also due to Euler):

$$B(\alpha, \beta) = \int_0^1 \mathrm{d}x\, x^{\alpha-1}(1-x)^{\beta-1} = \frac{\Gamma_E(\alpha)\Gamma_E(\beta)}{\Gamma_E(\alpha+\beta)}. \tag{E.15}$$

To obtain it, consider the product of two gamma functions:

$$\begin{aligned}\Gamma(\alpha)\Gamma(\beta) &= \int_0^\infty \mathrm{d}s \int_0^\infty \mathrm{d}t\, s^{\alpha-1} t^{\alpha-1}\, \mathrm{e}^{-(s+t)} \\ &= \int_0^\infty \mathrm{d}u \int_0^1 \mathrm{d}v\, u^{\alpha+\beta-1}\, v^{\alpha-1}(1-v)^{\beta-1}\, \mathrm{e}^{-u} \\ &= \Gamma_{\mathrm{E}}(\alpha+\beta)\, B(\alpha,\beta).\end{aligned} \tag{E.16}$$

This completes the proof of eq. (E.15).

E.3 Properties of the *N*-Dimensional Space

We will often have to deal with integrals in N-dimensional spaces, where N is very large. Since some geometrical properties of such spaces are far from intuitive, it is useful to have a closer look at them.

Let us first remark that the ratio between the volume of a spherical ball of radius $R - \delta R$ ($\delta R > 0$) and that of a ball of radius R is given by $(1 - \delta R/R)^N \approx \exp(-N\,\delta R/R)$. This ratio, however small δR is, approaches 0 as N increases—in other words, almost all of the ball's volume is found close to its surface.

The surface of a ball of radius R in N dimensions is given by $S_N(R) = S_N(1)\, R^{N-1}$. In order to calculate $S_N(1)$, we evaluate the N-dimensional Gaussian integral in polar coordinates:

$$G_N = \int \prod_{i=1}^N \mathrm{d}x_i\, \exp\left(-\sum_{i=1}^N x_i^2\right) = \int_0^\infty \mathrm{d}r\, R^{N-1}\, S_N(1)\, \exp(-R^2). \tag{E.17}$$

This integral is the product of N Gaussian integrals and is therefore equal to $\pi^{N/2}$. On the other hand, one has

$$\begin{aligned}G_N &= S_N(1) \int_0^\infty \mathrm{d}R\, R^{N-1}\, \exp\left(-R^2\right) \\ &= \frac{1}{2} S_N(1) \int_0^\infty \mathrm{d}t\, t^{N/2-1}\, \mathrm{e}^{-t} = \frac{1}{2} S_N(1)\, \Gamma_{\mathrm{E}}(N/2),\end{aligned} \tag{E.18}$$

where we use the definition of the gamma function. We therefore obtain

$$S_N(R) = \frac{2\pi^{N/2}}{\Gamma_{\mathrm{E}}(N/2)} R^{N-1}. \tag{E.19}$$

Let us now evaluate the projection of the ball onto one of its axes—in other words, the density of the points of the ball whose first coordinate lies, for example, between x and $x + \mathrm{d}x$. This quantity is given by

$$\rho(x) = \int \prod_{i=2}^N \mathrm{d}x_i\, \theta_{\mathrm{H}}\left(R^2 - x^2 - \sum_{i=2}^N x_i^2\right), \tag{E.20}$$

where $\theta_{\mathrm{H}}(x)$ is Heaviside's theta function:

$$\theta_{\mathrm{H}}(x) = \begin{cases} 1, & \text{if } x > 0; \\ 0, & \text{otherwise.} \end{cases} \tag{E.21}$$

The integral on the right-hand side is the volume of the ball of radius $\sqrt{R^2 - x^2}$ in $(N-1)$-dimensional space. We can therefore write

$$\begin{aligned} \rho(x) &= \frac{1}{N-1} S_{N-1}(1) \left(R^2 - x^2\right)^{(N-1)/2} \\ &\simeq \frac{S_{N-1}(1) R^{N-1}}{N-1} \exp\left[\frac{N-1}{2} \ln\left(1 - \frac{x^2}{R^2}\right)\right]. \end{aligned} \tag{E.22}$$

Because of the large factor $(N-1)$ in the exponential, this function becomes very small as soon as x moves away from zero. By expanding the exponent in a Taylor series around the origin, we obtain

$$\rho(x) \propto \exp\left[\frac{N-1}{2} \ln\left(1 - \frac{x^2}{R^2}\right)\right] \approx \exp\left[-\frac{(N-1)x^2}{2R^2}\right]. \tag{E.23}$$

The distribution is therefore a Gaussian of zero average and variance equal to $R^2/(N-1)$.

Appendix F Basics of Information Theory

In this appendix we present the essential tools of information theory. We define the Shannon entropy of a probability distribution, and two related quantities, namely, the Kullback-Leibler divergence (also called the relative entropy) and the mutual information. An excellent introduction to the subject is found in [Cove06].

F.1 Shannon Entropy

The **Shannon entropy** $H_S(p)$ of a probability distribution $p = (p_x)$ is the only measure of the inherent uncertainty of p that satisfies the following requirements:

1. $H_S(p)$ is a continuous function of the p_x. Therefore small changes in the p_xs entail only small changes of $H_S(p)$.
2. If $x \in \{1, \ldots, N\}$ and $p_x = 1/N$, $\forall x$, $\mathcal{I}(N) = H_S\left(\frac{1}{N}, \ldots, \frac{1}{N}\right)$ must be an increasing function of N. Indeed, we expect the uncertainty to increase with an increasing number of equally probable possibilities.
3. We require the uncertainty measure to be consistent, meaning that if we can evaluate it in more than one way, the results we obtain must agree.

To clarify the last requirement, suppose that $x \in \{1, \ldots, N\}$ with probabilities $p = (p_x)$. Gather the first k_1 alternatives in a box $y = 1$, and assign to it the probability $w_1 = p_1 + \cdots + p_{k_1}$. Then gather the next k_2 alternatives in a box $y = 2$, and assign to it the probability $w_2 = p_{k_1+1} + \cdots + p_{k_1+k_2}$. Proceed in the same way until the last alternatives are gathered in the box $y = K$. Thus we have either N alternatives with the probability distribution $p = (p_x)$, or K alternatives y with probability $w = (w_y)$, such that, if it is known that the system is in box y, the probability distribution of x in the box is given by $p_{x|y} = p_x/w_y$. We require therefore

$$H_S(p) = H_S(w) + \sum_y w_y H_S(p_{x|y}). \tag{F.1}$$

In particular, if the xs and the ys are all equally likely, and $N = K \cdot M$, we should have

$$\mathcal{I}(N) = \mathcal{I}(K) + \mathcal{I}(M). \tag{F.2}$$

Therefore, if the decomposition of N into prime factors is $q_1^{m_1} \cdots q_r^{m_r}$, we have

$$\mathcal{I}(N) = \sum_{\ell=1}^{r} m_\ell \mathcal{I}_{q_\ell}. \tag{F.3}$$

Requiring $\mathcal{I}(N)$ to be monotonically increasing in N implies that $\mathcal{I}(q)/\ln q = \text{const.}$ for any q prime. To see this, take any two different primes, s and t. Then, for arbitrarily large n, we can find a number m such that

$$\frac{m}{n} \leq \frac{\ln t}{\ln s} \leq \frac{m+1}{n}. \tag{F.4}$$

Since $\mathcal{I}(n)$ is monotonically increasing, we must have $\mathcal{I}(s^m) \leq \mathcal{I}(t^m) \leq \mathcal{I}(s^{m+1})$. We thus obtain

$$\frac{m}{n} \leq \frac{\mathcal{I}(t)}{\mathcal{I}(s)} \leq \frac{m+1}{n}. \tag{F.5}$$

From the relations (F.4) and (F.5) we obtain

$$\left| \frac{\ln t}{\ln s} - \frac{\mathcal{I}(t)}{\mathcal{I}(s)} \right| \leq \frac{1}{n}, \tag{F.6}$$

which can be recast as follows:

$$\left| \frac{\mathcal{I}(s)}{\ln s} - \frac{\mathcal{I}(t)}{\ln t} \right| \leq \frac{\mathcal{I}(s)}{n \ln t}. \tag{F.7}$$

Since the right-hand side can be made arbitrarily small, we obtain that $\mathcal{I}(s)/\ln s = k$ cannot depend on s.

We have therefore

$$\mathcal{I}(N) = k \ln N, \tag{F.8}$$

where $k > 0$ is an arbitrary constant. Since $H_S(p)$ is continuous, it will be fully defined once we are able to evaluate it for distributions $p = (p_x)$ with rational values of the form $p_x = n_x / \sum_{x'} n_{x'}$. To obtain its expression, we suppose that we have $N = \sum_x n_x$ alternatives, and we assign to each x n_x equally likely alternatives. We then have by eq. (F.2)

$$\mathcal{I}(N) = H_S(p) + \sum_x p_x \mathcal{I}(n_x). \tag{F.9}$$

By eq. (F.8) we then obtain, taking into account that $\sum_x p_x = 1$, the expression of the Shannon entropy:

$$H_S(p) = -k \left(\sum_x p_x \ln n_x - \ln N \right) = -k \sum_x p_x \ln p_x, \tag{F.10}$$

where k is a positive constant. In information-theoretical contexts the constant k is taken to be $1/\ln 2$, so that the entropy of the distribution of two equally likely alternatives is equal

to unity. This unit is called one **bit** of information. In thermodynamics, the natural choice is the Boltzmann constant k_B.

The **Kullback-Leibler divergence** (also called the **relative entropy**) between two different probability distributions $p = (p_x)$ and $q = (q_x)$ of the same variable x is defined by

$$D_{KL}(p\|q) = \sum_x p_x \ln \frac{p_x}{q_x}. \tag{F.11}$$

By means of the inequality

$$\ln x \leq x - 1, \tag{F.12}$$

we obtain

$$-D_{KL}(p\|q) = \sum_x p_x \ln \frac{q_x}{p_x} \leq \sum_x p_x \left(\frac{q_x}{p_x} - 1\right) = \sum_x q_x - \sum_x p_x = 0. \tag{F.13}$$

Therefore $D_{KL}(p\|q) \geq 0$, as stated in sec. 10.3, with equality holding only if $p_x = q_x$, $\forall x$. One has, in general, $D_{KL}(p\|q) \neq D_{KL}(q\|p)$. Indeed, if, for example, $q_x = 0$ for some values of x such that $p_x > 0$, one has $D_{KL}(p\|q) = \infty$, but $D_{KL}(q\|p)$ may be finite.

F.2 Mutual Information

The **mutual information** of two random variables, x and y, is a measure of the information shared by them. It is defined by

$$I(x:y) = \sum_{x,y} p_{x,y} \ln \frac{p_{x,y}}{p_x\, p_y}. \tag{F.14}$$

It is easy to see that it is equal to the Kullback-Leibler divergence between the joint distribution $p_{x,y}$ and the product of the marginal distributions $p_x = \sum_y p_{x,y}$ and $p_y = \sum_x p_{x,y}$. It is therefore nonnegative, and vanishes only if x and y are independent. It can also be expressed by

$$I(x:y) = \sum_{x,y} p_{x,y} \ln \frac{p_{x|y}}{p_x} = \sum_{x,y} p_{x,y} \ln \frac{p_{y|x}}{p_y}, \tag{F.15}$$

where $p_{x|y}$ is the distribution of x, given y. From eq. (F.14) we see that the mutual information is symmetrical. It is related to the **conditional entropy** $H_S(x|y)$ defined by

$$H_S(x|y) = -\sum_{x,y} p_{x,y} \ln p_{x|y}. \tag{F.16}$$

We have indeed

$$\begin{aligned} I(x:y) &= \sum_{x,y} p_{x,y} \ln \frac{p_{x|y}}{p_x} = -\sum_x p_x \ln p_x + \sum_{x,y} p_{x,y} \ln p_{x|y} \\ &= H_S(p_x) - H_S(x|y). \end{aligned} \tag{F.17}$$

We see therefore that $I(x:y)$ expresses the decrease in the uncertainty in x due to the information gleaned by y.

Appendix G Markov Chains

In the simulation with the Monte Carlo method, we introduce a fictitious dynamical process for the system under consideration, chosen so as to guarantee that the corresponding stationary distribution is the Boltzmann-Gibbs one. This process has the following characteristics:

- It takes place at discrete instants in time.
- At each such instant, the state of the system univocally determines the probabilities of its states at the next instant.

A dynamical process with these properties is called a **Markov chain**. In the following sections, we provide a general definition of Markov chain and prove (informally) some fundamental theorems. Unless stated otherwise, we will only consider processes with a finite number of states.

G.1 Definitions

Let us consider a system that can be, at each instant $t = 0, 1, 2, \ldots$, in one of N states. We denote by $X_t \in \{1, 2, \ldots, N\}$ the state of the system at the instant t. The sequence $(X_0, X_1, \ldots, X_t, \ldots)$ of the random variables X_t constitutes a random process, which is called a **Markov chain** if the probability distribution of X_{t+1} depends only on the value of X_t. In other words, given the state of the system at a given instant, the probability distribution of its states at the next instant is fully determined and does not depend on the previous history of the process. The dynamics of the system is therefore determined by the collection of all conditional probabilities of X_{t+1} given the values of X_t. They are called the **transition probabilities** of the process.

Let us denote the transition probability from x' to x at time t by $W_{xx'}(t)$—in other words, we have

$$W_{xx'}(t) = P(X_{t+1} = x | X_t = x'). \tag{G.1}$$

This definition does not rule out cases where the transition probabilities $W_{xx'}(t)$ depend on time. However, we only consider here Markov chains whose transition matrix $\mathbb{W} = (W_{xx'})$ is independent of t, known as **time-homogeneous** Markov chains.

Let us denote by $p_x(t)$ the probability that $X_t = x$, $x \in \{1, \ldots, N\}$. Then $\mathbb{W}$ and $p(t) = (p_x(t))$ obviously have the following properties:

1. $p_x(t) \geq 0$, $\forall x$.
2. $\sum_x p_x(t) = 1$, $\forall t$ (normalization of p).
3. $W_{xx'} \geq 0$, $\forall x, x'$.
4. $\sum_x W_{xx'} = 1$, $\forall x'$ (conservation of probability).
5. $p_x(t+1) = \sum_{x'} W_{xx'}\, p_{x'}(t)$, $\forall i$ (by the law of total probabilities).

A matrix $\mathbb{W}$ that satisfies properties 3 and 4 is called **stochastic**. It follows from property 5 that

$$p_x(t+n) = \sum_{x'} W^{(n)}_{xx'} p_{x'}(t), \qquad \forall t, n \geq 0, \tag{G.2}$$

where we denote by $W^{(n)}_{xx'}$ the matrix elements of the nth power of $\mathbb{W}$.

Taking advantage of property 4, we can write the evolution equation for $p(t)$ in the following form:

$$\Delta p_x(t+1) = p_x(t+1) - p_x(t) = \sum_{x' (\neq x)} \left[W_{xx'} p_{x'}(t) - W_{x'x} p_x(t) \right]. \tag{G.3}$$

The first term in brackets represents the inflow of probability to state x from different states, while the second one represents the outflow of probability from x to the other states. This equation is called the **master equation**.

G.2 Spectral Properties

Now let λ be an eigenvalue of $\mathbb{W}$. One has the following properties:

1. 1 is an eigenvalue.
2. $|\lambda| \leq 1$, $\forall \lambda$.
3. If $v^{(\lambda)} = \left(v^{(\lambda)}_x\right)$ is a right eigenvector that belongs to λ, then either $\lambda = 1$ or $\sum_x v^{(\lambda)}_x = 0$.

This can be proved as follows.

1. Consider the vector $u^{(1)}$ such that $u^{(1)}_x = 1$, $\forall x$. As a result of property 4, we have

$$1 = u^{(1)}_x = \sum_{x'} W_{xx'} = \sum_{x'} u^{(1)}_x W_{xx'}, \qquad \forall x. \tag{G.4}$$

Therefore $u^{(1)}$ is a left eigenvector of $\mathbb{W}$ that belongs to the eigenvalue 1.

2. Let $v^{(\lambda)} = \left(v_x^{(\lambda)}\right)$ be a right eigenvector of $\mathbb{W}$ that belongs to the eigenvalue λ. We then have

$$\sum_{x'} W_{xx'} \left|v_{x'}^{(\lambda)}\right| \geq \left|\sum_{x'} W_{xx'} v_{x'}^{(\lambda)}\right| = |\lambda| \left|v_x^{\lambda}\right|. \tag{G.5}$$

Taking the sum of this relation with respect to x and taking into account property 4, we obtain

$$\sum_{x} \left|v_x^{(\lambda)}\right| \geq |\lambda| \sum_{x} \left|v_x^{(\lambda)}\right|. \tag{G.6}$$

Therefore $|\lambda| \leq 1$. Moreover, if $|\lambda| = 1$, one must have

$$\sum_{x'} W_{xx'} \left|v_{x'}^{(\lambda)}\right| = \left|v_x^{(\lambda)}\right|, \qquad \forall x. \tag{G.7}$$

Thus $\left(\left|v^{(\lambda)}\right|_x\right)$ is a right eigenvector that belongs to the eigenvalue 1.

3. Let $v^{(\lambda)}$ be a right eigenvector of $\mathbb{W}$ that belongs to λ:

$$\sum_{x'} W_{xx'} v_{x'}^{(\lambda)} = \lambda v_x^{(\lambda)}. \tag{G.8}$$

Taking the sum of this equation with respect to x we obtain

$$\sum_{x'} v_{x'}^{(\lambda)} = \lambda \sum_{x'} v_{x'}^{(\lambda)}, \tag{G.9}$$

and therefore either $\lambda = 1$ or $\sum_{x'} v_{x'}^{(\lambda)} = 0$.

This completes the proof.

Let $v^{(1)} = \left(v_x^{(1)}\right)$ be a right eigenvector belonging to eigenvalue 1, satisfying $\sum_{x'} W_{xx'} v_{x'}^{(1)} = v_x^{(1)}$. Then, as we have seen, $\left(\left|v^{(1)}\right|_x\right)$ is a right eigenvector belonging to the eigenvalue 1. Let us define $p^{\mathrm{st}} = (p_x^{\mathrm{st}})$ by

$$p_x^{\mathrm{st}} = \frac{\left|v_x^{(1)}\right|}{\sum_{x'} \left|v_{x'}^{(1)}\right|}, \qquad \forall x. \tag{G.10}$$

In this way p^{st} represents a probability distribution, called a **stationary distribution**, since it is a constant solution to the master equation:

$$\sum_{x' (\neq x)} \left[W_{xx'} p_{x'}^{\mathrm{st}} - W_{x'x} p_x^{\mathrm{st}}\right] = 0. \tag{G.11}$$

G.3 Convergence to Stationarity

From now on, let us assume that the transition matrix $\mathbb{W} = (W_{xx'})$ is **ergodic**. This means that that there is an integer n such that the nth power $\mathbb{W}^n$ of $\mathbb{W}$ has all nonzero entries. In this situation, the eigenvalue equal to 1 is not degenerate, and all other eigenvalues are strictly smaller than 1 in magnitude. While it is possible to prove these results by

an indirect argument, we provide in the following a direct proof of the convergence to a stationary distribution for ergodic Markov chains.

By the results of sec. G.2, there exists a vector $p^{\mathrm{st}} = (p_x^{\mathrm{st}})$ such that

1. $p_x^{\mathrm{st}} \geq 0$, $\forall x$.
2. $\sum_x p_x^{\mathrm{st}} = 1$.
3. $\sum_{xx'} W_{xx'} p_{x'}^{\mathrm{st}} = p_x^{\mathrm{st}}$, $\forall x$.

We now wish to show that if $p(0) = (p_x(0))$ is a vector with non-negative entries that satisfies $\sum_i p_x(0) = 1$, one has

$$\lim_{t\to\infty} \sum_{xx'} W_{xx'}^{(t)} p_{x'}(0) = p_x^{\mathrm{st}}, \qquad \forall x, \tag{G.12}$$

where $W_{xx'}^{(t)}$ are the matrix elements of $\mathbb{W}^t$. In other words, starting from any initial probability distribution $p(0)$, the evolved probability distribution $p(t) = \mathbb{W}^t \cdot p(0)$ approaches the stationary distribution p^{st} as $t \to \infty$.

We can assume without losing generality that all entries of $\mathbb{W}$ are positive, by changing the transition matrix from $\mathbb{W}$ to $\mathbb{W}^n$ (we denote it again by $\mathbb{W}$ in the following). We then have $p_x^{\mathrm{st}} > 0$, $\forall x$. Indeed, if this were not the case, there would be at least one state x_0 such that $\sum_x W_{x_0 x'} p_{x'}^{\mathrm{st}} = 0$. But this contradicts the hypothesis that all entries of the matrix $\mathbb{W}$ are positive.

We obviously obtain the desired result (G.12) for any initial distribution $p(0)$ if we can prove that

$$\lim_{t\to\infty} W_{xx'}^{(t)} = p_x^{\mathrm{st}}, \qquad \forall x, x'. \tag{G.13}$$

We now have, for $t \geq 1$,

$$W_{xx'}^{(t)} = \sum_y W_{xy}^{(t-1)} W_{yx'} \geq \left(\min_y W_{xy}^{(t-1)} \right) \sum_y W_{yx'} = \min_y W_{xy}^{(t-1)}. \tag{G.14}$$

This equality holds also for the entry x' such that $W_{xx'}^{(t)} = \min_y W_{xy}^{(t)}$. Therefore

$$\min_{x'} W_{xx'}^{(t)} \geq \min_{x'} W_{xx'}^{(t-1)}. \tag{G.15}$$

One similarly obtains

$$\max_{x'} W_{xx'}^{(t)} \leq \max_{x'} W_{xx'}^{(t-1)}. \tag{G.16}$$

Thus we obtain two monotonic sequences, which are obviously bounded, the first by 1 above and the second by 0 below. As a consequence, they converge, and we have

$$\lim_{t\to\infty} \min_{x'} W_{xx'}^{(t)} = p_x^{(-)}, \qquad \lim_{t\to\infty} \max_{x'} W_{xx'}^{(t)} = p_x^{(+)}, \tag{G.17}$$

where $p_x^{(-)} \leq p_x^{(+)}$, $\forall x$.

Let us now show that $\lim_{t\to\infty} p_x^{(-)} = \lim_{t\to\infty} p_x^{(+)} = p_x^{\infty}$, $\forall x$. Consider the difference

$$W_{yx}^{(t)} - W_{yx'}^{(t)} = \sum_z W_{yz}^{(t-1)} (W_{zx} - W_{zx'}). \tag{G.18}$$

For each y, let $\Delta_{y,xx'}^{(+)} = W_{yx} - W_{yx'}$ if $W_{yx} > W_{yx'}$ and zero otherwise. Analogously, define $\Delta_{y,xx'}^{(-)} = -\left(W_{yx} - W_{yx'}\right)$ if $W_{yx} < W_{yx'}$ and zero otherwise. We have, by the condition $\sum_x W_{xx'} = 1$,

$$\sum_y \left(W_{yx} - W_{yx'}\right) = \sum_y \Delta_{y,xx'}^{(+)} - \sum_y \Delta_{y,xx'}^{(-)} = 0. \tag{G.19}$$

We obtain therefore

$$\sum_y \Delta_{y,xx'}^{(+)} = \sum_y \Delta_{y,xx'}^{(-)} = h_{xx'}. \tag{G.20}$$

Since all entries of $\mathbb{W}$ are positive, and assuming that not all of them are equal, we have

$$\sum_y \Delta_{y,xx'}^{(+)} < \sum_y W_{yx} = 1, \tag{G.21}$$

and therefore

$$0 \le h_{xx'} < 1. \tag{G.22}$$

Therefore, since the number of states is finite,

$$h = \max_{xx'} h_{xx'} < 1. \tag{G.23}$$

From eq. (G.18) we obtain

$$\begin{aligned}
\left|W_{xx'}^{(t)} - W_{yx'}^{(t)}\right| &= \left|\sum_z W_{xz}^{(t-1)} \left(W_{zx'} - W_{zy}\right)\right| \\
&= \left|\sum_z W_{xz}^{(t-1)} \left(\Delta_{z,x'y}^{(+)} - \Delta_{z,x'y}^{(-)}\right)\right| \\
&\le \left|\max_z W_{xz}^{(t-1)} \cdot \sum_z \Delta_{x,x'y}^{(+)} - \min_z W_{xz}^{(t-1)} \cdot \sum_z \Delta_{z,x'y}^{(-)}\right| \\
&\le h \left|\max_z W_{xz}^{(t-1)} - \min_z W_{xz}^{(t-1)}\right| \\
&\le h \max_{yz} \left|W_{xz}^{(t-1)} - W_{xy}^{(t-1)}\right|.
\end{aligned} \tag{G.24}$$

We have therefore, taking the max of this relation,

$$\max_{x'y} \left|W_{xx'}^{(t)} - W_{xy}^{(t)}\right| \le h \max_{x'y} \left|W_{xx'}^{(t-1)} - W_{xy}^{(t-1)}\right|, \tag{G.25}$$

where $h < 1$. Iterating t times, we obtain

$$\max_{x'y} \left|W_{xx'}^{(t)} - W_{xy}^{(t)}\right| \le h^t \max_{x'y} \left|W_{xx'} - W_{xy}\right|. \tag{G.26}$$

Since h^t converges exponentially to 0 for $t \to \infty$, we obtain

$$\lim_{t\to\infty} \left| W^{(t)}_{xx'} - W^{(t)}_{xy} \right| = 0, \qquad \forall x', y, \tag{G.27}$$

which implies

$$\lim_{t\to\infty} p^{(+)}_x(t) = \lim_{t\to\infty} p^{(-)}_x(t) = p^\infty_x. \tag{G.28}$$

Now p^∞ must be equal to p^{st}. If this were not the case, the vector $\alpha p^\infty + \beta p^{\mathrm{st}}$, with α and β arbitrary, would be an eigenvector of $\mathbb{W}$ belonging to the eigenvalue 1, and such would also be the case for the vector with entries $\left|\alpha p^\infty_x + \beta p^{\mathrm{st}}_x\right|$. We could then choose α and β such that one or more components (but not all) of this vector vanish. But this is not possible, since it contradicts the hypothesis that all entries of $\mathbb{W}$ are positive. This implies that the eigenvalue λ is nondegenerate and the convergence of $W^{(t)}_{xx'}$ to p^{st}_x implies that $|\lambda| \leq h < 1$ for all eigenvalues λ different from 1.

G.4 Non-ergodic Matrices

What happens if $\mathbb{W}$ is not ergodic? Here we discuss only the case where the diagonal elements of $\mathbb{W}$ are all positive. We say that two states x and x' are **connected** if it is possible to find a sequence $(x_0, x_1, \ldots, x_r)$ of states such that $x_0 = x$, $x_r = x'$, and $W_{x_{n+1}x_n} > 0$ for $n = 0, \ldots, r-1$. It is then easy to show that the relation "being connected" is reflexive, symmetric, and transitive, and is therefore an equivalence relation. Therefore the states split into equivalence classes $K^{(\alpha)}$, and the restriction $\mathbb{W}^{(\alpha)} = \left(W^{(\alpha)}_{xx'}\right)$, $x, x' \in K^{(\alpha)}$, to each equivalence class is ergodic, while all elements of $\mathbb{W}$ connecting states of two different classes vanish. Applying to each such restriction the method of sec. G.3, we find that there is a unique stationary distribution $p^{(\alpha)}$ defined in it. Therefore any combination of the form $\sum_\alpha c_\alpha p^{(\alpha)}$, where $c_\alpha \geq 0$, $\forall \alpha$, and $\sum_\alpha c_\alpha = 1$, is also a stationary distribution. Thus the eigenvalue 1 is m-fold degenerate, where m is the number of distinct equivalence classes $K^{(\alpha)}$. Physically, this corresponds to a situation where m Markov chains act in parallel and independently, each within the states of a single equivalence class, and the coefficients c_α represent the probability of being in class $K^{(\alpha)}$, which is conserved by the dynamics.

Appendix H Fundamental Physical Constants

The IUPAP-approved notation for uncertainty is of the form, e.g., $\ell = 10.50(15)$ m, where it is understood that the true value lies between 10.35 m and 10.65 m. The values reported here follow the 2019 redefinition of the SI base units, which fixed exact values for Planck's constant, Boltzmann's constant, Avogadro's number, the speed of light, and the elementary charge. The standard molar volume of the ideal gas, V_{mol}, is defined by IUPAC as the volume of one mole of ideal gas at the temperature of 273.15 K and a pressure of 100 kPa. Thus the given value is exact. The value of the standard gravitational acceleration g_0 was fixed at 9.806 65 m s^{-2} in the third General Conference on Weights and Measures in 1901.

Planck constant	h	$6.626\,070\,15 \cdot 10^{-34}$ J s (exact)
	$\hbar = h/(2\pi)$	$1.054\,571\,817\ldots 10^{-34}$ J s
Boltzmann constant	k_{B}	$1.380\,649 \cdot 10^{-23}$ J K^{-1} (exact)
Elementary charge	e	$1.602\,176\,634 \cdot 10^{-19}$ C (exact)
Avogadro number	N_{A}	$6.022\,140\,76 \cdot 10^{23}$ mol^{-1} (exact)
Speed of light	c	299 792 458 m s^{-1} (exact)
Vacuum permeability	μ_0	$1.256\,637\,062\,12(19) \cdot 10^{-6}$ N A^{-2}
Vacuum permittivity	ϵ_0	$8.854\,187\,812\,8(13) \cdot 10^{-12}$ F m^{-1}
Atomic mass constant	m_{u}	$1.660\,539\,066\,60(50) \cdot 10^{-27}$ kg
Bohr radius	a_0	$5.291\,772\,109\,03(80) \cdot 10^{-11}$ m
Gas constant	$R = N_{\text{A}}\, k_{\text{B}}$	8.314 462 618 153 24 J K^{-1} mol^{-1} (exact)
Standard molar volume	V_{mol}	$22.71\,095\,464\,148\,557\,506\,0 \cdot 10^{-3}$ m^3 (exact)
Faraday constant	$F = N_{\text{A}} e$	$9.648\,533\,212\,331\,001\,84 \cdot 10^{4}$ C mol^{-1} (exact)
Electron volt		$1.602\,176\,634 \cdot 10^{-19}$ J (exact)
		$k_{\text{B}} \cdot 1.160\,451\,812\ldots 10^{4}$ K (exact)
Von Klitzing constant	$R_{\text{K}} = h/e^2$	$2.581\,280\,745\ldots 10^{4}\ \Omega$ (exact)
Josephson constant	$K_{\text{J}} = 2e/h$	$4.835\,978\,484\ldots 10^{14}$ Hz/V (exact)
Fine structure constant	$1/\alpha = (2\epsilon_0 hc)/e^2$	137.035 999 084 (21)
Electron mass	m_{e}	$9.109\,383\,701\,5(28) \cdot 10^{-31}$ kg

Proton mass	m_p	$1.672\,621\,923\,69(51) \cdot 10^{-27}$ kg
Bohr magneton	$\mu_B = eh/(4\pi m_e)$	$9.274\,010\,0783(28) \cdot 10^{-24}$ J T^{-1}
Nuclear magneton	$\mu_N = eh/(4\pi m_p)$	$5.050\,783\,699\,(31) \cdot 10^{-27}$ J T^{-1}
Electron gyromagnetic ratio	g_e	2.002 319 304 361 18(27)
Electron magnetic moment	$\mu_e = -(1/2)g_e\mu_B$	$-9.284\,764\,7043(28) \cdot 10^{-24}$ J T^{-1}
Proton magnetic moment	μ_p	$1.410\,606\,797\,36(60) \cdot 10^{-26}$ J T^{-1}
Gravitational constant	G	$6.674\,30(15) \cdot 10^{-11}$ J m kg^{-2}
Standard gravity acceleration	g_0	9.806 65 m s^{-2} (exact)

Bibliography

[Abe73] R. Abe and S. Hikami. Critical exponents and scaling relations in $1/n$ expansion. *Progress of Theoretical Physics*, 49:442–452, 1973.

[Ahle80] G. Ahlers. Critical phenomena at low temperature. *Rev. Mod. Phys.*, 52:489–503, 1976.

[Alle87] M. P. Allen and D. J. Tildesley. *Computer Simulations of Liquids*. Oxford University Press, Oxford, UK, 1987.

[Alme78] J. R. L. de Almeida and D. J. Thouless. Stability of the Sherrington-Kirkpatrick solution of a spin-glass model. *J. Phys. A: Math. Gen.*, 11:983–990, 1978.

[AlsN76] J. Als-Nielsen. Neutron scattering and spatial correlation near the critical point. In C. Domb and M. S. Green, editors, *Phase Transitions and Critical Phenomena*, volume 5a, pages 87–164. Academic Press, 1976.

[Amit05] D. J. Amit and V. Martín-Mayor. *Field Theory, the Renormalization Group, and Critical Phenomena*. World Scientific, Singapore, 3rd edition, 2005.

[Amit85] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Spin-glass models of neural networks. *Phys. Rev. A*, 32:1007–1018, 1985.

[Amit89] D. J. Amit. *Modelling Brain Function: The World of Attractor Neural Networks*. Cambridge University Press, Cambridge, UK, 1989.

[Ande80] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72:2384–2393, 1980.

[Appe77] K. Appel and W. Haken. Every planar map is four colorable. *Illinois Journal of Mathematics*, 21:439–567, 1977.

[Araj61] S. Arajs and R. V. Colvin. Paramagnetism of polycrystalline gadolinium, terbium and dysprosium metals. *J. Appl. Phys.*, 32 (suppl.):S336–S337, 1961.

[Ashc76] N. Ashcroft and N. D. Mermin. *Solid State Physics*. Holt, Rinehart and Winston, New York, 1976.

[Bald21] M. Baldovin, S. Iubini, R. Livi, and A. Vulpiani. Statistical mechanics of systems with negative temperature. *Physics Reports*, 923:1–50, 2021.

[Bali91] R. Balian. *From Microphysics to Macrophysics: Methods and Applications of Statistical Physics*. Springer, Berlin/New York, 1991.

[Bark76] J. A. Barker and D. W. Henderson. What is "liquid"? Understanding the states of matter. *Rev. Mod. Phys.*, 48:587–671, 1976.

[Barr03] J.-L. Barrat and J.-P. Hansen. *Basic Concepts for Simple and Complex Liquids*. Cambridge University Press, Cambridge, UK, 2003.

[Batt20] G. Battimelli, G. Ciccotti, and P. Greco. *Computer Meets Theoretical Physics: The New Frontier of Molecular Simulation*. Springer, Berlin, 2020.

[Benz82] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. Stochastic resonance in climatic change. *Tellus*, 34:10–16, 1982.

[Bere71] V. L. Berezinskii. Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group. I. Classical systems. *Sov. Phys. JETP*, 32:493–500, 1971.

[Berg55] P. G. Bergmann and J. L. Lebowitz. New approach to nonequilibrium processes. *Phys. Rev.*, 99:578–587, 1955.

[Berr97] M. V. Berry and A. K. Geim. Of flying frogs and Levitrons. *European Journal of Physics*, 18:307–313, 1997.

[Bind19] K. Binder and D. W. Heermann. *Monte-Carlo Simulations in Statistical Physics*. Graduate Texts in Physics. Springer Nature Switzerland AG, Cham, 6th edition, 2019.

[Bind86] K. Binder. *Monte Carlo Methods in Statistical Mechanics*. Springer, Berlin, 1986.

[Binn92] J. J. Binney, N. J. Dowrick, A. J. Fisher, and M. Newman. *The Theory of Critical Phenomena*. Clarendon Press, Oxford, 1992.

[Bish78] D. J. Bishop and J. D. Reppy. Study of the superfluid transition in two-dimensional ^{4}He films. *Physical Review Letters*, 40:1727–1730, 1978.

[Blac73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[Bolt95] L. Boltzmann. *Vorlesungen über Gastheorie*. Barth, Leipzig, 1895–1898. (2 vols.).

[Born12a] M. Born and T. von Kármán. On fluctuations in spatial grids. *Physikalische Zeitschrift*, 13:297–309, 1912.

[Born12b] M. Born and T. von Kármán. On the theory of specific heat. *Physikalische Zeitschrift*, 14:15–19, 1912.

[Bose24] S. N. Bose. Plancks Gesetz und Lichtquantenhypothese. *Zeitschrift für Physik*, 26:178–181, 1924.

[Bott21] L. Böttcher and H. J. Herrmann. *Computational Statistical Physics*. Cambridge University Press, Cambridge, UK, 2021.

[Bouc97] J.-P. Bouchaud and M. Mézard. Universality classes for extreme-value statistics. *J. Phys. A: Math. Gen.*, 30:7997–8016, 1997.

[Brez10] E. Brézin. *Introduction to Statistical Field Theory*. Cambridge University Press, Cambridge, UK, 2010.

[Broa57] S. R. Broadbent and J. M. Hammersley. Percolation processes: I. Crystals and mazes. *Proceedings of the Cambridge Philosophical Society*, 53:629–641, 1957.

[Brow20] R. G. Brown. Robert G. Brown's General Tools Page. https://webhome.phy.duke.edu/∼rgb/General/dieharder.php, 2020.

[Brus76] S. G. Brush. *The Kind of Motion We Call Heat*. North-Holland, Amsterdam, 1976.

[Call85] H. B. Callen. *Thermodynamics and an Introduction to Thermostatistics*. J. Wiley & Sons, New York, 1985.

[Cara09] C. Carathéodory. Untersuchungen über die Grundlagen der Thermodynamik. *Mathematische Annalen*, 67:355–386, 1909. English translation in J. Kestin (ed.). *The Second Law of Thermodynamics*, pages 229–256. Dowden, Hutchison and Ross, Stroudsburg PA, 1976.

[Card96] J. Cardy. *Scaling and Renormalization in Statistical Physics*. Cambridge University Press, Cambridge, UK, 1996.

[Carn69] N. F. Carnahan and K. E. Starling. Equation of state for nonattracting rigid spheres. *J. Chem. Phys.*, 51:635–636, 1969.

[Chai95] P. M. Chaikin and T. C. Lubensky. *Principles of Condensed Matter Physics*. Cambridge University Press, Cambridge, UK, 1995.

[Chan31] S. Chandrasekhar. The maximum mass of ideal white dwarfs. *Astrophysical Journal*, 74:81–82, 1931.

[Chan87] D. Chandler. *Introduction to Modern Statistical Mechanics*. Oxford University Press, Oxford, UK, 1987.

[Cloi90] J. des Cloiseaux and J.-F. Jannink. *Polymers in Solution: Their Modelling and Structure*. Clarendon Press, Oxford, 1990.

[Clus34] K. Clusius and E. Bartholomé. Die Rotationswärmen der Moleküle HD und D_2 und der Kernspin des D-Atoms. *Zeitschrift für Elektrochemie und angewandte physikalische Chemie*, 40:524–529, 1934.

[Cole73] S. Coleman and E. Weinberg. Radiative corrections as the origin of spontaneous symmetry breaking. *Phys. Rev. B*, 7:1888–1910, 1973.

[Coni80] A. Coniglio and W. Klein. Clusters and Ising critical droplets: A renormalisation group approach. *J. Phys. A: Math. Gen.*, 13:2775–2780, 1980.

[Conn71] D. L. Connelly, J. S. Loomis, and D. E. Mapother. Specific heat of nickel near the Curie temperature. *Phys. Rev. B*, 3:924–934, 1971.

[Cove06] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. J. Wiley & Sons, Hoboken, NJ, 2006.

[Curi95] P. Curie. Les propriétés magnétiques des corps à diverses températures. *Journal de Physique*, 4:197–263, 1895.

[Deby12] P. Debye. Zur Theorie der spezifischen Wärme. *Ann. Phys. (Leipzig)*, 344:789–839, 1912.

[Deby23] P. Debye and E. Hückel. The theory of electrolytes. I. Lowering of freezing point and related phenomena. *Physikalische Zeitschrift*, 24:185–206, 1923.

[Derr81] B. Derrida. Random-energy model: An exactly solvable model of disordered systems. *Phys. Rev. B*, 24:2613–2626, 1981.

[DeSo53] W. DeSorbo. Specific heat of diamond at low temperatures. *J. Chem. Phys.*, 21:876–880, 1953.

[DiCa69] C. Di Castro and G. Jona-Lasinio. On the microscopic foundation of scaling laws. *Physics Letters*, 29A:322–323, 1969.

[Dick65] R. H. Dicke, P. J. E. Peebles, P. J. Roll, and D. T. Wilkinson. Cosmic black-body radiation. *Astrophysical Journal Letters*, 142:414–419, 1965.

[Doi13] M. Doi. *Soft Matter Physics*. Oxford University Press, Oxford, 2013.

[Doi96] M. Doi. *Introduction to Polymer Physics*. Clarendon Press, Oxford, 1996.

[Domb71] C. Domb and M. S. Green. *Phase Transitions and Critical Phenomena*. Academic Press, London, 1971–2001. 20 vols. Later edited by C. Domb and J. L. Lebowitz.

[Dots01] V. Dotsenko. *Introduction to the Replica Theory of Disordered Statistical Systems*. Cambridge University Press, Cambridge, UK, 2001.

[Dunn75] A. G. Dunn, J. W. Essam, and J. M. Loveluck. Scaling theory for the pair-connectedness in percolation models. *J. Phys. C: Solid State Phys.*, 8:743–750, 1975.

[Earl05] D. J. Earl and M. W. Deem. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.*, 7:3910–3916, 2005.

[Edwa75] S. F. Edwards and P. W. Anderson. Theory of spin glasses. *J. Phys. F: Metal. Phys.*, 5:965–974, 1975.

[Ehre33] P. Ehrenfest. Phasenumwandlungen im ueblichen und erweiterten Sinn, classifiziert nach dem entsprechenden Singularitaeten des thermodynamischen Potentiales. *Verhandlingen der Koninklijke Akademie van Wetenschappen (Amsterdam) 3*, 36:153–157, 1933.

[Eins05] A. Einstein. Die von der Molekularkinetische Theorie von Wärme geforderte Bewegung von in den ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys. (Leipzig)*, 17:549–560, 1905. Reprinted in A. Einstein. *Investigations on the Theory of Brownian Movement*, pages 1–18. Dover, New York, 1956.

[Eins07] A. Einstein. Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Wärme. *Ann. Phys. (Leipzig)*, 22:180–190, 1907.

[Eins07a] A. Einstein. Theoretische Bemerkungen über die Brownsche Bewegung. *Zeit. f. Elektrochemie*, 13:41–42, 1907. Reprinted in A. Einstein. *Investigations on the Theory of Brownian Movement*, pages 63–67. Dover, New York, 1956.

[Eins10] A. Einstein. Theorie der Opaleszenz von homogenen Flüssigkeiten und Flüssigkeitsgemischen in der Nähe des kritischen Zustandes. *Ann. Phys. (Leipzig)*, 33:1275–1298, 1910.

[Eins25] A. Einstein. Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, (I):3–14, 1925.

[Eins56] A. Einstein. *Investigations on the Theory of Brownian Movement*. Dover, New York, 1956. Edited with notes by R. Fürth.

[Ferm55] E. Fermi, J. Pasta, and S. Ulam. Studies of nonlinear problems. Technical Report LA-1940, Los Alamos National Laboratory, 1955. Reprinted in: E. Fermi. *Collected Papers*, Vol. 2, No. 266, pages 978–988. The University of Chicago Press, Chicago, 1965.

[Feyn72] R. P. Feynman. *Statistical Mechanics: A Set of Lectures*. Benjamin, Reading, MA, 1972.

[FIRAS] FIRAS Collaboration. FIRAS project data set. https://lambda.gsfc.nasa.gov/data/cobe/firas/monopole_spec/firas_monopole_spec_v1.txt, 2005.

[Fisc77] K. H. Fischer. Spin glasses. *Physica B+C*, 86-88:813–819, 1977.

[Fisc91] K. H. Fischer and J. A. Hertz. *Spin Glasses*. Cambridge University Press, Cambridge, UK, 1991.

[Fish83] M. E. Fisher. Scaling, universality, and renormalization group theory. In F. J. Hahne, editor, *Critical Phenomena, Proceedings of the Summer School Held at the University of Stellenbosch, South Africa, January 18–29, 1982*, pages 1–139. Springer, Berlin, 1983.

[Flor53] P. J. Flory. *Principles of Polymer Chemistry*. Cornell University Press, Ithaca, NY, 1953.

[Fors75] D. Forster. *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions*. Benjamin, Reading, MA, 1975.

[Fort72] C. M. Fortuin and P. W. Kasteleyn. On the random-cluster model: I. Introduction and relation to other models. *Physica (Utrecht)*, 57:536–564, 1972.

[Fren01] D. Frenkel and B. Smit. *Understanding Molecular Simulation*, volume 1 of *Computational Science Series*. Academic Press, 2nd edition, 2001.

[Gall95] G. Gallavotti and E. G. D. Cohen. Dynamical ensembles in nonequilibrium statistical mechanics. *Physical Review Letters*, 74:2694–2694, 1995.

[Gall99] G. Gallavotti. *Statistical Mechanics: A Short Treatise*. Springer, Berlin, 1999.

[Gamm98] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni. Stochastic resonance. *Reviews of Modern Physics*, 70:223–287, 1998.

[Gard09] C. Gardiner. *Stochastic Methods*. Springer, Berlin, 4th edition, 2009.

[Genn79] P.-G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca, NY, 1979.

[Genn93] P.-G. de Gennes and J. Prost. *The Physics of Liquid Crystals*. Clarendon Press, Oxford, 2nd edition, 1993.

[Ghir98] S. Ghirlanda and F. Guerra. General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A: Math. Gen.*, 46:9149–9155, 1998.

[Gibb02] J. W. Gibbs. *Elementary Principles in Statistical Mechanics*. Ch. Scribner's Sons, New York, 1902.

[Gibb73] J. W. Gibbs. Graphical methods in the thermodynamics of fluids. *Trans. Conn. Acad. Arts and Sci.*, 2:309–342, 1873.

[Gibb75] J. W. Gibbs. On the equilibrium of heterogeneous substances. *Trans. Conn. Acad. Arts and Sci.*, III:108–248, 343–524, 1875-76.

[Ginz60] V. L. Ginzburg. Some remarks on phase transitions of the 2nd kind and the microscopic theory of ferroelectric materials. *Soviet Physics-Solid State*, 2:1824, 1960.

[Gold20] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì. Gibbs and Boltzmann entropy in classical and quantum mechanics. In V. Allori, editor, *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, pages 519–581. World Scientific, Singapore, 2020.

[Gold92] N. Goldenfeld. *Lectures of Phase Transitions and the Renormalization Group*. CRC Press, Boca Raton, FL, 1992.

[Gouy88] L. G. Gouy. Note sur le mouvement Brownien. *J. Phys. Theor. Appl.*, 7:561–564, 1888. Available online at https://hal.science/jpa-00238904/document.

[Guer02] F. Guerra and F. L. Toninelli. The thermodynamic limit in mean field spin glass models. *Commun. Math. Phys.*, 230:71–79, 2002.

[Guer03] F. Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Commun. Math. Phys.*, 233:1–12, 2003.

[Guer81] F. Guerra. Structural aspects of stochastic mechanics and stochastic field theory. *Physics Reports*, 77:263–312, 1981.

[Gugg45] E. A. Guggenheim. The principle of corresponding states. *J. Chem. Phys.*, 13:253– 261, 1945.

[Hans13] J.-P. Hansen and I. R. McDonald. *Theory of Simple Liquids with Applications to Soft Matter*. Academic Press, Amsterdam-New York, 4th edition, 2013.

[Harr74a] A. B. Harris and T. C. Lubensky. Renormalization-group approach to the critical behavior of random-spin models. *Phys. Rev. Lett.*, 33:1540–1543, 1974.

[Harr74b] A. B. Harris. Effect of random defects on the critical behaviour of Ising models. *J. Phys. C: Solid State Phys.*, 7:1671–1692, 1974.

[Hart05] A. Hartmann and M. Weigt. *Phase Transitions in Combinatorial Optimization Problems*. Wiley-VCH, Weinheim, 2005.

[Hirs54] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird. *Molecular Theory of Gases and Liquids*. John Wiley & Sons, New York, 1954.

[Hohe64] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.

[Hoog92] P. J. Hoogerbrugge and J. M. V. A. Koelman. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhysics Letters*, 19:155–160, 1992.

[Hoov85] W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, 1985.

[Hopf82] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA*, 79:2554–2558, 1982.

[Hove54] L. van Hove. Time-dependent correlations between spins and neutron scattering in ferromagnetic crystals. *Physical Review*, 95:1374–1384, 1954.

[Huan87] K. Huang. *Statistical Mechanics*. J. Wiley & Sons, New York, 2nd edition, 1987.

[Hups42] J. C. Hupse. The magnetic susceptibilities of copper potassium sulphate hexahydrate. *Physica*, IX:633–640, 1942.

[Imry75] Y. Imry and S.-K. Ma. Random-field instability of the ordered state of continuous symmetry. *Phys. Rev. Lett.*, pages 1399–1401, 1975.

[Isin25] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31:253–258, 1925.

[Isra91] J. Israelachvili. *Intermolecular & Surface Forces*. Academic Press, London, 2nd edition, 1991.

[Ito44] K. Ito. Stochastic integral. *Proc. Imp. Acad. Tokyo*, 20:519, 1944.

[Jarz11] C. Jarzynski. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. *Annual Reviews of Condensed Matter Physics*, 2:329–351, 2011.

[Jayn03] E. T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, UK, 2003. Edited by G. L. Bretthost.

[Jayn80] E. T. Jaynes. The minimum entropy production principle. *Ann. Rev. Phys. Chem.*, 31:579–601, 1980.

[Jayn92] E. T. Jaynes. The Gibbs paradox. In C. R. Smith, G. J. Erickson, and P. O. Neudorfer, editors, *Maximum Entropy and Bayesian Methods*, pages 1–22. Kluwer Academic Publishers, Dordrecht, 1992.

[Jose77] J. V. José, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson. Renormalization, vortices, and symmetry-breaking perturbations in the two-dimensional planar model. *Physical Review B*, 16:1217–1241, 1977.

[Kac52] M. Kac and J.-C. Ward. A combinatorial solution of the two-dimensional Ising model. *Phys. Rev.*, 88:1332–1337, 1952.

[Kada66] L. P. Kadanoff. Scaling laws for Ising models near T_c. *Physics*, 2:263–272, 1966.

[Kada67] L. P. Kadanoff, W. Götze, D. Hamblen, R. Hecht, E. A. S. Lewis, V. V. Palciauskas, M. Rayl, J. Swift, D. Aspnes, and J. Kane. Static phenomena near critical points: Theory and experiment. *Rev. Mod. Phys.*, 39:395–431, 1967.

[Kada76] L. P. Kadanoff. Notes on Migdal's recursion formula. *Ann. Phys. (N.Y.)*, 100:359–394, 1976.

[Kapp31] E. Kappler. Versuche zur Messung der Avogadro-Loschmidtschen Zahl aus der Brownschen Bewegung einer Drehwaage. *Ann. Phys. (Leipzig)*, 403:233–256, 1931.

[Kirk83] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.

[Klei01] H. Kleinert and V. Schulte-Frohlinde. *Critical Properties of ϕ^4 Theories*. World Scientific, Singapore, 2001.

[Knut97] D. E. Knuth. *Seminumerical Algorithms*, volume 2 of *The Art of Computer Programming*. Addison-Wesley, Reading, MA, 1997.

[Kofk02] D. A. Kofke. On the acceptance probability of replica-exchange Monte Carlo trials. *The Journal of Chemical Physics*, 117:6911–6914, 2002.

[Kost73] J. M. Kosterlitz and D. J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C: Solid State Physics*, 6:1181–1203, 1973.

[Kram40] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.

[Kram41] H. A. Kramers and G. H. Wannier. Statistics of the two-dimensional ferromagnet. *Phys. Rev.*, 60:252–262, 1941.

[Krau06] W. Krauth. *Stiatistical Mechanics: Algorithms and Computations*. Oxford University Press, Oxford, UK, 2006.

[Krau21] W. Krauth. Event-chain Monte Carlo: Foundations, applications, and prospects. *Frontiers in Physics*, 9:663457, 2021.

[Kriv08] H. Krivine and J. Treiner. *La Physique Statistique en Exercices*. Vuibert, Paris, 2008.

[Kubo85] R. Kubo, M. Toda, and N. Hashitsume. *Statistical Physics II, Nonequilibrium Statistical Mechanics*. Springer, Berlin, 1965.

[Kurc98] J. Kurchan. Fluctuation theorem for stochastic dynamics. *Journal of Physics A: Mathematical and General*, 31:3719–3729, 1998.

[Land30] L. D. Landau. Diamagnetismus der metalle. *Zeitschrift für Physik*, 64:629–637, 1930.

[Land80] L. D. Landau and E. M. Lifshitz. *Statistical Physics, Part I*. Pergamon Press, Oxford, UK, 3rd edition, 1980. Edited by E. M. Lifshitz and L. P. Pitayevskii.

[Lang08] P. Langevin. Sur la théorie du mouvement Brownien. *C. R. Acad. Sci. Paris*, 146:530–533, 1908.

[Lang18] I. Langmuir. The adsorption of gases on plane surfaces of glass, mica and platinum. *J. Am. Chem. Soc.*, 40:1361–1403, 1918.

[Lebo93] J. L. Lebowitz. Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy. *Physica A*, 194:1–27, 1993.

[Lebo99] J. L. Lebowitz and H. Spohn. A Gallavotti-Cohen-type symmetry in the large deviation function for stochastic dynamics. *Journal of Statistical Physics*, 95:333–365, 1999.

[Lee52] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. *Phys. Rev.*, 87:410–419, 1952.

[Lieb98] E. H. Lieb and J. Yngvason. A guide to entropy and the second law of thermodynamics. *Notices of the American Mathematical Society*, 45:571–581, 1998.

[Lifs64] S. Lifson. Partition function of linear chain molecules. *J. Chem. Phys.*, 40:3705–3710, 1964.

[Lim90] Y.-K. Lim. *Problems and Solutions in Statistical Mechanics*. World Scientific, 1990.

[Lube79] T. C. Lubensky. Thermal and geometrical critical phenomena in random systems. In R. Balian, R. Maynard, and G. Toulouse, editors, *Ill-Condensed Matter, Les Houches Session XXXI, 1978*, pages 405–475. North-Holland, Amsterdam, 1979.

[Ma73] S.-K. Ma. Critical exponents above T_c to O($1/n$). *Physical Review A*, 7:2172–2187, 1973.

[Ma76] S.-K. Ma. *Modern Theory of Critical Phenomena*. W. A. Benjamin, Inc., Reading, MA, 1976.

[Ma85] S.-K. Ma. *Statistical Mechanics*. World Scientific, Singapore, 1985.

[Mand67] B. Mandelbrot. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science*, 156:636–638, 1967.

[Mand77] B. Mandelbrot. *Fractals: Form, Chance, and Dimension*. W. H. Freeman and Co., San Francisco, 1977.

[Mant99] R. N. Mantegna and H. E. Stanley. *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, Cambridge, UK, 1999.

[Marc59] P. M. Marcus and H. J. McFee. Velocity distributions in potassium molecular beams. In I. Estermann, editor, *Recent Research in Molecular Beams*, pages 43ff. Academic Press, New York, 1959.

[Mari92] E. Marinari and G. Parisi. Simulated tempering: A new Monte Carlo scheme. *Europhysics Letters (EPL)*, 19:451–458, 1992.

[Mars03] G. Marsaglia. Xorshift RNGs. *Journal of Statistical Software*, 8:1–6, 2003.

[Mars95] G. Marsaglia. The Marsaglia Random Number CDROM including the Diehard Battery of tests of randomness. Florida State University, 1995.

[Mart01] O. C. Martin, R. Monasson, and R. Zecchina. Statistical mechanics methods and phase transitions in optimization problems. *Theoretical Computer Science*, 265:3–67, 2001.

[Mart94] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein. Constant pressure molecular dynamics algorithms. *Journal of Chemical Physics*, 101:4177–4189, 1994.

[Mart95] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein. Explicit reversible integrators for extended systems dynamics. *Molecular Physics*, 87:1117–1157, 1995.

[Matt76] D. C. Mattis. Solvable spin systems with random interactions. *Physics Letters A*, 56:421–422, 1976.

[Maxw60] J. C. Maxwell. Illustrations of the dynamical theory of gases. Part 1. On the motion and collisions of perfectly elastic spheres. *Phil. Mag.*, XIX:19–32, 1860.

[Maxw75] J. C. Maxwell. On the dynamical evidence of the molecular constitution of bodies. *Nature*, 11:357–359, 1875.

[Maye40] J. E. Mayer and M. G. Mayer. *Statistical Mechanics*. Wiley, New York, 1940.

[Mazo02] R. M. Mazo. *Brownian Motion: Fluctuations, Dynamics and Applications*. Oxford University Press, Oxford, 2002.

[McCu43] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biol.*, 5:115–133, 1943.

[Merm65] N. D. Mermin. Thermal properties of the inhomogeneous electron gas. *Phys. Rev.*, 137:A1441–A1443, 1965.

[Merm66] N. D. Mermin and H. Wagner. Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic Heisenberg models. *Physical Review Letters*, 17:1133–1136, 1966.

[Meza09] M. Mézard and A. Montanari. *Information, Physics, and Computation*. Oxford University Press, Oxford, 2009.

[Meza15] M. Mézard. Cavity method: Message-passing from a physics perspective. In F. Krzakala, F. Ricci-Tersenghi, L. Zdeborova, R. Zecchina, E. W. Tramel, and L. F. Cugliandolo, editors, *Statistical Physics, Optimization, Inference, and Message-Passing Algorithms: Lecture Notes of the Les Houches School of Physics: Special Issue, October 2013*. Oxford University Press, 2015.

[Meza87] M. Mézard, G. Parisi, and M. A. Virasoro. *Spin Glass Theory and Beyond*. World Scientific, Singapore, 1987.

[Migd75] A. A. Migdal. Phase transitions in gauge and spin-lattice systems. *Zh. Eksp. Teor. Fiz. (USSR)*, 69:1457–1465, 1975. Translated in *Soviet Physics-JETP* 42:743–746, 1976.

[Milch86] A. Milchev, D. W. Heermann, and K. Binder. Finite-size scaling analysis of the ϕ^4 field theory on the square lattice. *Journal of Statistical Physics*, 44:749–784, 1986.

[Mori58] T. Morita. Theory of classical fluids: Hyper-Netted Chain approximation, I: Formulation for a one-component system. *Progress of Theoretical Physics*, 20:920–938, 1958.

[Mori65] H. Mori. Transport, collective motion, and Brownian motion. *Prog. Theor. Phys.*, 33:423–455, 1965.

[Moss99] S. Moss de Oliveira, P. M. C. de Oliveira, and D. Stauffer. *Evolution, Money, War, and Computers*. Teubner, Stuttgart, 1999.

[Naka58] S. Nakajima. On quantum theory of transport phenomena. *Prog. Theor. Phys.*, 20:948–959, 1958.

[Naue08] M. Nauenberg. Edmund C. Stoner and the discovery of the maximum mass of white dwarfs. *Journal for the History of Astronomy*, 39:297–312, 2008.

[Nels83] D. R. Nelson. Defect-mediated phase transitions. In C. Domb and J. L. Lebowitz, editors, *Phase Transitions and Critical Phenomena*, volume 7, pages 1–65. Academic Press, New York, 1983.

[Niem76] T. H. Niemeyer and J. M. J. van Leeuwen. Renormalization for Ising-like spin systems. In C. Domb and M. S. Green, editors, *Phase Transitions and Critical Phenomena*, volume 6, pages 425–505. Academic Press, New York, 1976.

[Nose84] S. Nosé. A unified formulation of the constant temperature molecular-dynamics methods. *Journal of Chemical Physics*, 81:511–519, 1984.

[Onsa31a] L. Onsager. Reciprocal relations in irreversible processes I. *Phys. Rev.*, 37:405–426, 1931.

[Onsa31b] L. Onsager. Reciprocal relations in irreversible processes II. *Phys. Rev.*, 38:2265–2279, 1931.

[Onsa44] L. Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65:117–149, 1944.

[Onsa52] L. Onsager. Interpretation of the de Haas-van Alphen effect. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 43:1006–1008, 1952.

[Orns14] L. S. Ornstein and F. Zernike. The accidental deviations of density and the opalescence in the critical point. *Proc. Acad. Sci. Amsterdam*, 17:793–806, 1914.

[Orns20] L. S. Ornstein. The Brownian movement. *Proc. Amst. Acad. Sci.*, 21:96, 1920.

[Pari09] G. Parisi. The mean-field theory of spin glasses: The heuristic replica approach and recent rigorous results. *Lett. Math. Phys.*, 88:255–269, 2009.

[Pari79] G. Parisi. The order parameter for spin glasses: A function on the interval 0–1. *J. Phys. A: Math. Gen.*, 13:1101–1112, 1979.

[Path11] R. K. Pathria and P. D. Beale. *Statistical Mechanics*. Elsevier, Amsterdam, 3rd edition, 2011.

[Peie34] R. Peierls. Remarks on transition temperatures. *Helv. Phys. Acta*, 7:81–83, 1934.

[Peli21] L. Peliti and S. Pigolotti. *Stochastic Thermodynamics: An Introduction*. Princeton University Press, Princeton, 2021.

[Penz65] A. A. Penzias and R. W. Wilson. A measurement of excess antenna temperature at 4080 Mc/s. *Astrophysical Journal Letters*, 142:419–421, 1965.

[Perc58] J. K. Percus and G. J. Yevick. Analysis of classical statistical mechanics by means of collective coordinates. *Physical Review*, 110:1–13, 1958.

[Perr08] J. Perrin. La loi de Stokes et le mouvement Brownien. *C. R. Acad. Sci. Paris*, 147:475–476, 1908.

[Perr10] J. Perrin. *Brownian Movement and Molecular Reality*. Taylor & Francis, London, 1910. Translated by F. Soddy.

[Perr16] J. Perrin. *Les Atomes*. Alcan, Paris, 1913. Translated in: J. Perrin, *Atoms*, Van Nostrand, New York, 1916.

[Pinc70] M. Pincus. A Monte-Carlo method for the approximate solution of certain types of constrained optimization problems. *Journal of the Operations Research Society of America*, 18:967–1235, 1970.

[Pipp57] A. B. Pippard. *The Elements of Classical Thermodynamics*. Cambridge University Press, Cambridge, UK, 1957.

[Plan01] M. Planck. Über das Gesetz der Energieverteilung im Normalspectrum. *Ann. Phys. (Leipzig)*, 4:553–563, 1901.

[Plis06] M. Plischke and B. Birgerson. *Equilibrium Statistical Mechanics*. World Scientific, Singapore, 3rd edition, 2006.

[Pool92] R. Pool. The third branch of science debuts. *Science*, 256:44–47, 1992.

[Pres07] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge (UK), 3rd edition, 2007.

[Prig47] I. Prigogine. *Etude thermodynamique des phénoménes irréversibles*. Desoer, Liège, 1947.

[Prig61] I. Prigogine. *Thermodynamics of Irreversible Processes*. Interscience, New York, 1961.

[Proh64] E. Prohofsky and J. Krumhansl. Second-sound propagation in dielectric solids. *Physical Review*, 133:A1403–A1410, 1964.

[Purc51] E. M. Purcell and R. V. Pound. A nuclear spin system at negative temperature. *Physical Review*, 81:279–280, 1951.

[Ramm86] R. Rammal, G. Toulouse, and M. A. Virasoro. Ultrametricity for physicists. *Rev. Mod. Phys.*, 58:765–788, 1986.

[Rams56] N. F. Ramsey. Thermodynamics and statistical mechanics at negative absolute temperatures. *Physical Review*, 103(1):20, 1956.

[Rath05] N. Rathore, M. Chopra, and J. J. de Pablo. Optimal allocation of replicas in parallel tempering simulations. *The Journal of Chemical Physics*, 122:024111, 2005.

[Reic80] L. E. Reichl. *A Modern Course in Statistical Physics*. University of Texas Press, Austin, 1980.

[Rich61] L. F. Richardson. The problem of contiguity: An appendix to statistics of deadly quarrels. *General Systems Yearbook*, page 139–187, 1961.

[Robe57] L. M. Roberts. The atomic heats of calcium, strontium and barium between 1.5° and 20°K. *Proceedings of the Physical Society. Section B*, 70:738–743, 1957.

[Rowl82a] J. S. Rowlinson and F. L. Swinton. *Liquids and Liquid Mixtures*, chapter 3. The critical state, pages 59–85. Butterworth Scientific, London, 3rd edition, 1982.

[Rowl82b] J. S. Rowlinson and F. L. Swinton. *Liquids and Liquid Mixtures*, chapter 4. The thermodynamics of liquid mixtures, pages 86–131. Butterworth Scientific, London, 3rd edition, 1982.

[Ruel69] D. Ruelle. *Statistical Mechanics: Rigorous Results*. Benjamin, New York, 1969.

[Sabe15] A. A. Saberi. Recent advances in percolation theory and its applications. *Physics Reports*, 578:1–32, 2015.

[Sack11] O. Sackur. On the kinetic explanation of Nernst's heat theory. *Ann. Phys. (Leipzig)*, 34:455–468, 1911.

[Safr94] S. A. Safran. *Statistical Thermodynamics of Surfaces, Interfaces, and Membranes*. Frontiers in Physics. Addison-Wesley, Reading, MA, 1994.

[Sant21] K. P. Santo and A. V. Neimark. Dissipative particle dynamics simulations in colloid and interface science: A review. *Advances in Colloid and Interface Science*, page 102545, 2021.

[Sarm79] G. Sarma. Conformation des polymères en solution. In R. Balian, R. Maynard, and G. Toulouse, editors, *Ill-Condensed Matter, Les Houches XXXI, 1978*, pages 539–552. North-Holland, Amsterdam, 1979.

[Schn76] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern Physics*, 48:571–585, 1976.

[Schu64] T. D. Schulz, D. C. Mattis, and E. H. Lieb. Two-dimensional Ising model as a soluble problem of many fermions. *Rev. Mod. Phys.*, 36(856-871), 1964.

[Seif12] U. Seifert. Stochastic thermodynamics, fluctuation theorems, and molecular machines. *Reports on Progress in Physics*, 75:126001, 2012.

[Seif18] U. Seifert. Stochastic thermodynamics: From principles to the cost of precision. *Physica A: Statistical Mechanics and Its Applications*, 504:176–191, 2018.

[Seki10] K. Sekimoto. *Stochastic Energetics*. Springer, Berlin, 2010.

[Seth06] J. P. Sethna. *Statistical Mechanics: Entropy, Order Parameters, and Complexity*. Oxford University Press, Oxford, 2006.

[Shan48] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423 & 623–656, 1948. Reprinted and expanded in: Shannon, C. E. and Weaver, W. *The Mathematical Theory of Communication*, University of Illinois Press, Urbana and Chigago, 1949.

[Sher75] D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Phys. Rev. Lett.*, 35:1792–1796, 1975.

[Smol06] M. von Smoluchowski. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Ann. Phys. (Leipzig)*, 326:756–780, 1906.

[Stan71] H. E. Stanley. *Introduction to Phase Transitions and Critical Phenomena*. Oxford University Press, Oxford, 1971.

[Stauf92] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. Taylor & Francis, London, 2nd edition, 1992.

[Ston29] E. C. Stoner. V. The limiting density in white dwarf stars. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 7:63–70, 1929.

[Suth05] W. Sutherland. A dynamical theory of diffusion for non-electrolytes and the molecular mass of albumin. *Philosophical Magazine*, 9:781–785, 1905.

[Sved06] T. Svedberg. Über die Eigenbewegung der Teilchen in kolloidalen Lösungen. *Zeit. für Elektrochemie*, 12:853, 1906.

[Swen86] R. H. Swendsen and J.-S. Wang. Replica Monte Carlo simulation of spin glasses. *Physical Review Letters*, 57:2607–2609, 1986.

[Swen87] R. H. Swendsen and J.-S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Physical Review Letters*, 58:86–88, 1987.

[Tala06] M. Talagrand. The Parisi formula. *Ann. Math.*, 163:221–263, 2006.

[Tala11a] M. Talagrand. *Mean Field Models for Spin Glasses, Volume 1: Basic Examples*, volume 54 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge*. Springer, Berlin, Heidelberg, 2011.

[Tala11b] M. Talagrand. *Mean Field Models for Spin Glasses, Volume II: Advanced Replica-Symmetry and Low Temperature*, volume 55 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge*. Springer, Berlin, Heidelberg, 2011.

[Taub14] U. Täuber. *Critical Dynamics*. Cambridge University Press, Cambridge, UK, 2014.

[Tetr14] H. M. Tetrode. Theoretical determination of the entropy constant of gases and liquids. *Proc. Sect. Sci. Koninklijke Nederlandse Akad. Wet., Ser. B*, 17:1167–1184, 1914.

[Tisz66] L. Tisza. *Generalized Thermodynamics*. The MIT Press, Cambridge MA, 1966.

[Tole87] J.-C. Tolédano and P. Tolédano. *The Landau Theory of Phase Transitions: Application to Structural, Incommensurate, Magnetic, and Liquid Crystal Systems*. World Scientific, Singapore, 1987.

[Tolm38] R. C. Tolman. *The Principles of Statistical Mechanics*. Oxford University Press, Oxford, 1938. Reprinted Dover, New York, 1979.

[Toul75] G. Toulouse and P. Pfeuty. Critère de Ginzburg et dimensionnalité caractéristique pour le probleme de la localisation dans les systèmes désordonnés. *C. R. Acad. Sci. Paris, Série B*, 280:33–35, 1975.

[Vdov64] N. V. Vdovichenko. A calculation of the partition function for a plane dipole lattice. *Zh. Eksp. Teor. Fiz. (USSR)*, 47:715–720, 1964. Translated in *Soviet Physics-JETP* 20:477–479 (1964).

[Vdov65] N. V. Vdovichenko. Spontaneous magnetization of a plane dipole lattice. *Zh. Eksp. Teor. Fiz. (USSR)*, 48:526–530, 1965. Translated in *Soviet Physics-JETP* 21:350–352 (1965).

[Verl67] L. Verlet. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159:98–103, 1967.

[Waal73] J. J. van der Waals. *Over de continuiteit van den gas- en vloeistoftoestand*. PhD thesis, Universiteit Leiden, 1873. Translated in: van der Waals, J. D., *On the Continuity of the Gaseous and Liquid States*, J. S. Rowlinson (Ed.). Dover, Minneola NY, 1988.

[Wann66] G. H. Wannier. *Statistical Physics*. J. Wiley & Sons, New York, 1966.

[Weis07] P. Weiss. L'hypothèse du champ moléculaire et la propriété ferromagnétique. *Journal de physique pure et appliquée*, 4ème série, 6:429–445, 1907.

[Weis26] P. Weiss and R. Forrer. Aimantation et phénomène magnétocalorique du nickel. *Ann. Phys. (Paris)*, 10:153–213, 1926.

[Wert64] M. S. Wertheim. Analytic solution of the Percus-Yevick equation. *Journal of Mathematical Physics*, 5:643–651, 1964.

[Wido63] B. Widom. Some topics in the theory of fluids. *Journal of Chemical Physics*, 39:2808–2812, 1963.

[Wien30] N. Wiener. Generalized harmonic analysis. *Acta Mathematica*, 55:117–258, 1930.

[Wils72] K. G. Wilson and J. Kogut. The renormalization group and the ϵ-expansion. *Physics Reports*, 12:75–200, 1972.

[Wils79] K. G. Wilson. Problems in physics with many scales of length. *Scientific American*, 241(2):140–157, 1979.

[Wolf89] U. Wolff. Collective Monte Carlo updating for spin systems. *Phys. Rev. Lett.*, 62:361–364, 1989.

[Yang52] C. N. Yang and T. D. Lee. Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Phys. Rev.*, 87:404–409, 1952.

[Zinn96] J. Zinn-Justin. *Quantum Field Theory and Critical Phenomena*. Clarendon Press, Oxford, 3rd edition, 1996.

[Zwan01] R. Zwanzig. *Nonequilibrium Statistical Mechanics*. Oxford University Press, Oxford, 2001.

[Zwan54] R. Zwanzig. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.*, 22:1420–1426, 1954.

[Zwan60] R. Zwanzig. Ensemble method in the theory of irreversibility. *J. Chem. Phys.*, 33:1338–1341, 1960.

Author Index

Abe, R. 245
Aharony, A. 441, 466
Ahlers, G. 205
Allen, M. P. 337
Als-Nielsen, J. 205
Amit, D. J. 200, 231, 253, 261, 465, 467
Andersen, H. C. 314
Anderson, P. W. 446, 457
Appel, K. 305
Arajs, S. 191
Ashcroft, N. 123
Aspnes, D. 206
Avogadro, A. 2

Baldovin, M. 80
Balian, R. 101
Barker, J. A. 275, 294
Barrat, J.-L. 304
Bartholomé, E. 151
Beale, P. D. 154
Benzi, R. 362
Berezinskii, V. L. 259
Bergmann, P. G. 396
Berry, M. V. 129
Binder, K. 306, 338
Binney, J. J. 261
Bird, R. B. 284
Birgersen, B. 154, 261
Bishop, D. J. 260
Black, F. 385
Bogolyubov, N. N. 273
Boltzmann, L. 4, 9, 59, 101, 262
Born, M. 111, 273
Bose, S. N. 141
Bouchaud, J.-P. 450
Broadbent, S. R. 426
Brown, R. G. 327
Brush, S. G. 10
Brézin, E. 261
Böttcher, L. 307, 338

Callen, H. B. xxi, 11, 58, 349
Carathéodory, C. 25
Cardy, J. 261
Carnahan, N. F. 289
Cauchy, A. L. 413
Chaikin, P. M. 215, 259
Chandler, D. 218, 221, 384
Chandrasekhar, S. 138
Chopra, M. 336
Clusius, K. 151
Cohen, E. G. D. 399
Coleman, S. 253
Colvin, R. V. 191
Coniglio, A. 436
Connelly, D. L. 186
Cover, T. M. 504
Crooks, G. 400
Curie, P. 88, 191
Curtiss, C. F. 284

de Almeida, J. R. L. 461
de Gennes, P. G. 466
de Oliveira, P. M. C. 338
de Pablo, J. J. 336
Debye, P. 111, 299
Deem, M. W. 336
Derrida, B. 448

des Cloiseaux, J. 425
DeSorbo, W. 112
Di Castro, C. 215
Dicke, R. H. 109
Doi, M. 304, 425, 466
Domb, C. 215
Dotsenko, V. 467
Dowrick, N. J. 261
Dunn, A. G. 427

Earl, D. J. 336
Edwards, S. F. 446, 457
Ehrenfest, P. 185
Einstein, A. 3, 105, 141, 279, 340, 346, 347, 349, 385
Essam, J. W. 427

Fermi, E. 308
Feynman, R. P. 12
FIRAS Collaboration, 110
Fischer, K. H. 446, 467
Fisher, A. J. 261
Fisher, M. E. 155
Flannery, B. P. 325–327, 334
Flory, P. J. 422
Forster, D. 384
Fortuin, C. M. 434
Frenkel, D. 337

Gallavotti, G. 101, 399
Gammaitoni, L. 362
Gardiner, C. 341, 385
Geim, A. K. 129
Gelatt Jr, C. D. 334
Ghirlanda, S. 462
Gibbs, J. W. xxi, 11, 76, 96, 101, 151
Ginzburg, V. L. 201
Goldenfeld, N. 231, 261
Goldstein, S. 101
Gouy, L. G. 340, 346
Green, H. S. 273
Green, M. S. 215
Guerra, F. 341, 462
Guggenheim, E. A. 160
Gutfreund, H. 465
Götze, W. 206

Haken, W. 305
Halperin, B. 261
Hamblen, D. 206
Hamilton, W. R. 60
Hammersley, J. M. 426
Hansen, J.-P. 281, 289, 304
Harris, A. B. 444
Hartmann, A. 467
Hashitsume, N. 365, 384
Hecht, R. 206
Heermann, D. W. 306, 338
Helmholtz, H. 391
Henderson, D. W. 275, 294
Herrmann, H. J. 307, 338
Hertz, J. A. 467
Hikami, S. 245
Hirschfelder, J. O. 284
Hohenberg, P. 277
Hoogerbrugge, P. J. 318
Hoover, W. G. 315, 316
Hopfield, J. J. 462, 464
Huang, K. 10, 154
Hupse, J. C. 88
Hänggi, P. 362
Hückel, E. 299

Imry, Y. 445
Ising, E. 3, 78, 165
Israelachvili, J. 304
Ito, K. 341
Iubini, S. 80

Jannink, J.-F. 425
Jarzynski, C. 399, 415
Jaynes, E. T. 75, 101, 382
Jensen, J. 481
Jona-Lasinio, G. 215
José, J. V. 259
Jung, P. 362

Kac, M. 210
Kadanoff, L. P. xxii, 206, 215, 224, 259, 438
Kane, J. 206
Kappler, E. 103
Kasteleyn, P. W. 434
Kirkpatrick, S. 259, 334, 457
Kirkwood, J. G. 273, 274
Klein, M. L. 316, 318
Klein, W. 436
Kleinert, H. 244, 261
Knuth, D. E. 326
Koelman, J. M. V. A. 318
Kofke, D. A. 336
Kogut, J. 215, 231
Kohn, W. 277
Kosterlitz, J. M. 255, 259, 261
Kramers, H. A. 179, 359
Krauth, W. 322, 338, 437
Krivine, H. 154
Krumhansl, J. 117
Kubo, R. 365, 384
Kullback, S. 391
Kurchan, J. 399

Lagrange, J.-L. 484
Landau, L. D. xxi, 101, 131, 171, 210, 215
Langevin, P. 340, 342
Lebowitz, J. L. 101, 215, 396, 399
Lee, T. D. 165
Legendre, A.-M. 482
Leibler, R. 391
Lenz, W. 165
Lewis, E. A. S. 206
Lieb, E. H. 24, 26, 210, 215
Lifshitz, E. M. 101, 171, 210, 215
Lifson, S. 177
Lim, Y.-K. 154
Livi, R. 80
Loomis, J. S. 186
Loveluck, J. M. 427
Lubensky, T. C. 215, 259, 433, 444

Ma, S.-K. xxi, xxii, 9, 12, 101, 245, 337, 445
Mandelbrot, B. 346, 439
Mantegna, R. N. 385
Mapother, D. E. 186
Marchesoni, F. 362
Marcus, P. M. 8
Marinari, E. 336
Marsaglia, G. 327
Martin, O. C. 467
Martín-Mayor, V. 201, 231, 253, 261
Martyna, G. J. 316, 318
Massieu, F. 86
Mattis, D. C. 210, 215, 448
Maxwell, J. C. 3, 157, 159
Mayer, J. E. 281
Mayer, M. G. 281
Mazo, R. M. 384
McCulloch, W. S. 462
McDonald, I. R. 281, 289, 304
McFee, H. J. 8
Mermin, N. D. 123, 135, 255, 277
Migdal, A. A. 438
Milchev, A. 331
Monasson, R. 467
Montanari, A. 467
Mori, H. 365
Morita, T. 291
Moss de Oliveira, S. 338
Mézard, M. 450, 462, 467

Nakajima, S. 365
Nauenberg, M. 136
Neimark, A. V. 318
Nelson, D. R. 259–261
Newman, M. E. J. 261
Niemeyer, T. H. 229, 231
Nosé, S. 315

Onsager, L. 135, 179, 210, 372, 378
Ornstein, L. S. 189, 340

Palciauskas, V. V. 206
Parisi, G. 336, 362, 458, 461, 462, 467
Pasta, J. 308
Pathria, R. K. 154
Pauli, W. 3
Peebles, P. J. E. 109
Peierls, R. 173
Peliti, L. 415
Penzias, A. A. 109
Percus, J. K. 292
Perrin, J. 339, 340, 345, 352
Pfeuty, P. 254
Pigolotti, S. 415
Pincus, M. 334
Pippard, A. B. 58
Pitts, W. 462
Planck, M. 109
Plischke, M. 154, 261
Pool, R. 306
Pound, R. V. 80
Pres07 334
Press, W. H. 325–327, 334
Prigogine, I. 382
Prohofsky, E. 117
Prost, J. 304
Purcell, E. M. 80

Rammal, R. 462
Ramsey, N. F. 80
Rathore, N. 336
Rayl, M. 206
Reichl, L. E. 384
Reppy, J. D. 260
Richardson, L. F. 346
Roberts, L. M. 126
Roll, P. J. 109
Rowlinson, J. S. 206
Ruelle, D. 288

Saberi, A. A. 427
Sackur, O. 75
Safran, S. A. 304
Santo, K. P. 318
Sarma, G. 423
Schnakenberg, J. 396
Scholes, M. 385
Schulte-Frohlinde, V. 244, 261
Schulz, T. D. 210, 215
Schwarz, H. A. 413
Seifert, U. 415
Sekimoto, K. 414
Sethna, J. P. 215

Shannon, C. E. 96, 391
Sherrington, D. 457
Smit, B. 337
Sommerfeld, A. 124
Sompolinsky, H. 465
Spohn, H. 399
Stanley, H. E. 215, 385
Starling, K. E. 289
Stauffer, D. 338, 441, 466
Stoner, E. C. 136
Sutera, A. 362
Sutherland, W. 343
Svedberg, T. 340
Swendsen, R. H. 336, 436
Swift, J. 206
Swinton, F. L. 206

Talagrand, M. 462
Tetrode, H. M. 75
Teukolsky, S. A. 325–327, 334
Thomas, J. A. 504
Thouless, D. J. 255, 259, 261, 461
Tildesley, D. J. 337
Tisza, L. xxi, 11, 58
Tobias, D. J. 316, 318
Toda, M. 365, 384
Tolédano, J.-C. 215
Tolédano, P. 215
Tolman, R. C. 10
Toninelli, F. L. 462
Toulouse, G. 254, 462
Treiner, J. 154
Tsingou, M. 308
Tuckerman, M. E. 316, 318
Tumulka, R. 101
Täuber, U. 368

Ulam, S. 308

van der Waals, J. J. 157
van Hove, L. 368
van Leeuwen, J. M. J. 229, 231
Vdovichenko, N. V. 210, 214
Vecchi, M. P. 334
Verlet, L. 311
Vetterling, W. T. 325–327, 334
Virasoro, M. A. 462, 467
von Kármán, T. 111
von Meyer, J. 47
von Smoluchowski, M. 340
Vulpiani, A. 80, 362

Wagner, H. 135, 255
Wang, J.-S. 336, 436
Wannier, G. H. 179
Ward, J.-C. 210
Weaver, W. 96
Weigt, M. 467
Weinberg, E. 253
Weiss, P. 190
Wertheim, M. S. 292
Widom, B. 333, 402
Wiener, C. 340
Wiener, N. 341
Wilkinson, D. T. 109
Wilson, K. G. 215, 231
Wilson, R. W. 109
Wolff, U. 436

Yang, C. N. 165
Yevick, G. J. 292
Yngvason, J. 24, 26
Young, A. P. 261
Yvon, J. 273

Zanghì, N. 101
Zecchina, R. 467
Zee, A. xx
Zernike, F. 189
Zinn-Justin, J. 206
Zwanzig, R. 293, 365, 384

Subject Index

Page numbers in italics refer to tables and figures.

absolute temperature, 18, 20–22
action potential, 462
additive variable, 13. *See also* extensive variable
adiabatic, compressibility, 47
adiabatically equivalent, 25
adiabatic atmosphere, 47
adiabatic entropy production, 402
adiabatic processes, law of, 47
adiabatic transformations, 24–26
adiabatic walls, 14; mobile, 18–19, 23
adsorption, 143–46
affinities, 378; fluxes and, 378–80
almost certain/almost impossible/almost never/almost surely, 468
amphiphilic molecules, 301–4
analyticity, 374, 474; of trial free energy, 194–95
angular quantum number, 150
aniline-cyclohexane, 161
anisotropy: cubic, 252–53; quadratic, 246–47
annealed average, 443
annealed disorder, 416
annealing, 443
antiferromagnet, 171, 188
antivortex, 256, 257
Arrhenius law, 360
articulation point of diagrams, 286, 287
associative memory, 464; Hopfield model of, 462–66
asymptotic series, 500
asynchronous updating, 464
atmosphere, thermal gradient of, 47–48
atomic gases: internal degrees of freedom of, 146–48. *See also* ideal gas
atomic mass constant, 513
Atoms (Perrin), 339
attraction basin, 228
attractive tail, 163, 264
availability, 197, 364; generalized Brownian motion and, 366; in ϕ^4 model, 237; Taylor expansion of, 198
average of random variable, 470, 471
Avogadro's number, 2, 87, 365, 513
axon, 462, 463

B, magnetic induction vector, 42, 43
backward protocol, 396
backward trajectory, 396–97
basin of attraction, of fixed point, 228, 247
basis vectors, 497
BBGKY hierarchy, 273–75
Bethe lattice, 429; percolation on, 428–32
Big Bang, 109
binary mixtures: critical exponents of, 205; phase coexistence of, 161–62, 279
binomial distribution, 469
bit, 506
black body, 3
black body radiation: photons and, 106–11; Planck's law for, 106–9, 110
block transformation, 215–18
Bohr effect, 145
Bohr magneton, 514
Bohr radius, 513
Bohr-van Leeuwen theorem, 80
Boltzmann constant, 6, 45, 62, 87, 352, 513
Boltzmann factor, 82, 99
Boltzmann-Gibbs distribution, 416
bond percolation, renormalization and, 437–39
bond percolation problem, 426

Born-Oppenheimer approximation, 148, 152
Bose-Einstein condensation, 142. *See also* Einstein condensation
Bose factor, 108, 109, 115, 121, 141
Bose gas, equation of state of, 142
Bose statistics, 140–41; variational derivation of, 140–41
boson gas, 2, 117–22; Einstein condensation of, 141–43
bosons, 2, 118
Bragg elastic scattering, 270
bridge, diagrams, 290–91
Brillouin zone, 497; first, 123, 202, 233, 234; on lattice, 496–98; renormalization models of, 233, 234, 239
Brownian motion, 339–42; in *d* dimensions, 344; generalized, 365–68; of harmonic oscillator, 375; phenomenon of, 339n1; and stochastic dynamics, 387
Brownian particle: behavior of, in bistable potential, 361, 362; correlation functions, 354–56; experimental trajectory of, 345; Fokker-Planck equation and, 353; fractal properties of trajectories of, 345–47; mobility of, in fluid, 401; simulated trajectory of, 346; Smoluchowski equation of, 347–52

calculation (computer), 305
canonical ensemble, 81–83; definition of, 82; energy fluctuations in, 83–84; ideal gas in, 84–85; magnetization and, 80
canonical sampling, proof of, 318–19
Cantor set, 439, 440
carbon, coexisting phases of, 54
carbon dioxide: critical point of, 155; phase diagram of, 157
Carnahan-Starling formula, 289, 293
Carnot's theorem, 21, 159
Casimir effect, 107
Cauchy-Schwarz inequality, 413, 475
Cayley tree, 429
central limit theorem, 474–75; fluctuating force and, 340, 349; Monte Carlo method and, 320
central problem of thermodynamics, 14–15
Chandrasekhar limit, 138
characteristic function, 473–74
Chebyshev's inequality, 472
chemical equilibrium, 48–49; in gases, 154
chemical kinetics, 362–65; and collective variables, 364; and law of mass action, 362–63; and rate constants, 363; and rate equations, 363; reaction chain and, 364–65; reaction coordinate and, 364
chemical potential, 19, 23–24; measurement of, 39–40
Chinese philosophers' problem, 178–79
Clausius-Clapeyron equation, 51–52
cluster approximations, 187
clustering property, 172
clusters, 426; in percolation problem, 416, 426–28; size distribution of, 430–32
coexistence, of several phases, 53–54
coexistence curve, 52–53; Ising model and, 185; and lattice gas, 162–64; for liquids in vicinity of critical temperature, 160; liquid-vapor, 155–57; and phase diagram, 156. *See also* phase coexistence
coherence length, 177; block transformation and, 215–17; classical exponents and, 204; correlation function and, 208; definition of, 208; and finite-size scaling, 331; Kadanoff transformation and, 218, 220; in Ornstein-Zernike theory, 199–200; quadratic anisotropy and, 247; scaling laws and, 330, 341
coil-globule transition, 418
coincidence method, 337
collective variable, 364
colloidal suspensions, 301–4
comparable states, 25
complex fluids, 262
compressibility: adiabatic, 14, 47, 114; equation of state for, 269; isothermal, 47, 160, 269, 273, 279
computer. *See* numerical simulation
concavity, 15, 479
condensation temperature, 142
conditional entropy, 506
conditional expectation, 472
conditional probability, 471
configuration integral, 91, 163, 263
conjugate jump rates, 403
conjugate variable, 30
connection, 426, 512
consistency, 25; thermodynamic, 389–90
consolution point, 162
constants of the motion: classical, 68; quantum mechanical, 69–70
constitutive relation, 354
constrained extremum, 16, 485
convexity, 15, 479–81, 484
convolution, 495
coordination number: of Bethe lattice, 430; of Ising model, 165; of mean-field theory, 183
copying information, 410–12
correlation(s), 475–76
correlation function(s): critical exponents and, 195–96, 245; critical Hamiltonian, 226; crossover and, 247–51; definition of, 267; of densities, 267; direct, 273; dynamics of, 354–57; in Einstein theory of fluctuations, 198, 199; of Gaussian model, 236–37; Ginzburg criterion and, 200, 247; Kubo formula and, 358; Lorentzian, 356; and mean-field theory of Ising model, 188–92; Ornstein-Zernike form and, 236, 421; quadratic anisotropy and, 247; velocity, 343
correlation length, 190. *See also* coherence length

correlation matrix, 476
cosine coefficients, 493; series, 493
coverage dimension, 347, 347n2
covolume, 157
critical behavior, Einstein theory of fluctuations and, 196–99
critical behavior of Ising model: correlation function and, 188; critical exponents and, 195–96; duality and, 179–82; exponents for, 195–96; Landau theory and, 192–95; Peierls's argument, 172–76; specific heat and, 184–86; spontaneous magnetization and, 170, 184, 200, 213, 329
critical crossover, 247–51
critical density, 155, 156
critical exponents, 195–96; for block transformation, 215–18; classical, 196; crossover, 247; and finite lattice method, 229–31; at first order in ε, 240–44; in limit $n \to \infty$, 201–4; 1/n expansion of, 245; in ϕ^4 model, 237–40; of physical systems, 205; theoretical estimates of, for different models, 206; universality and scaling of, 204–9
critical Hamiltonian, 226
critical isotherm, 245
critical micellar concentration (cmc), 302
critical opalescence, 155
critical phenomena, linear polymers and, 423–25
critical point, 54–55; of carbon dioxide, 155; liquid-gas coexistence and, 155–56
critical pressure, 156
critical slowing down, 368
critical surface, 226
critical temperature, 54, 156
Crooks relation, 400
crossover, 247; critical, 247–51
crossover exponent, 247
cubic anisotropy, 252–53; fixed point of, 252
cumulant expansion, 238
cumulant generating function, 454
Curie point in ferromagnets, 55
Curie's law, 88
Curie temperature, 186
Curie-Weiss law, 190, 193
cutoff, 111, 233, 256

d'Alembert's equation, of sound waves, 114
damping coefficient, 340, 349, 354
de Broglie relation, 270
de Broglie thermal wavelength: classic fluids and, 302; definition of, 84; for degenerate boson gas, 120, 122, 141; for electrons at 300K, 123; ideal gas criterion and, 122; ideal gas partition function and, 91; for lattice gas, 163; phase transitions and, 166; for simple fluid, 263
Debye frequency, 111
Debye function, 112
Debye model, 111, 112
Debye screen length, 299
Debye temperature, 111
Debye wave vector, 111
decimation, 218; for free energy in one-dimensional Ising model, 221; in one-dimensional Ising model, 218–20
degenerate, 122, 123, 141
degree of polymerization, 418
degrees of freedom: internal, of atomic gases, 146–48; internal, of molecular gases, 148–51; three-dimensional, 151; vibrational and rotational, 149
de Haas-van Alphen effect, 133–35
delta distribution, 470
delta function. *See* Dirac delta
demagnetization factor, 43
dendrites (neuron), 462, 463
densities: continuously varying, 51; definition of, 30; discontinuously varying, 51; equations of state and, 44–45; intensive variables vs., 30–33, 44–45; pair distribution as, 266; in phase coexistence, 50. *See also* reduced densities
density functional approach, 276–80; and approximations to F[ρ], 277–79; fluid-fluid interfaces and, 279–80; and theory, 264
density matrix/density operator, 92
density of states, in metals, 123, 124, 126
detailed balance, 322, 388
diamagnetism, 80; classical theory of, 131; and De Haas-van Alphen effect, 133–35; and Landau, 131–33; Langevin, 129–31
Dieharder, 327
differential form, 488
diffusion coefficient, 343, 357–58, 401; definition of, 344; Einstein relation and, 342–44
diffusion current, 349
diffusion equation, 348, 354
diffusion process, 352; Fokker-Planck equation and, 352–53
dimensionality, 166, 209, 253, 268, 358, 445
dimensions, lower and upper critical, 253–55
Dirac delta "function," 6, 477–78, 498
direct correlation function, 273, 300
disorder: annealed, 416; quenched, 416
dissipative particle dynamics (DPD), 318
doping, 416
driving, 387
duality: and Ising model, 179–82; and three-dimensional Ising model, 182
dual lattice, Ising model and, 180
Dulong-Petit law, 105
dwells, 387
dynamics: affinities and fluxes and, 378–80; Brownian motion and, 339–42; chemical kinetics and, 362–65; correlation functions and, 354–57; fluctuation-dissipation theorem and, 372–76;

dynamics (*continued*)
generalized Brownian motion and, 365–68; Kubo formula and sum rules and, 357–58; metastable states and stochastic resonance and, 358–62; nonequilibrium steady states and, 380–82; Onsager reciprocity relations and, 376–78; response functions and, 368–72

Edwards-Anderson model, 446, 457
Edwards-Anderson order parameter, 446
Einstein condensation, 3, 141–43; in harmonic potential, 143; order parameter and, 171; of zero-mass particles, 143
Einstein relation, 343, 350; diffusion coefficient and, 342–44
Einstein temperature, 105, 106
Einstein theory of fluctuations, 196–99, 200, 208, 365
elasticity: elementary model of, 89; theory of, 57–58
electron: Fermi energy of, 136; mass, 513; in metals, 123–27
electron gyromagnetic ratio, 514
electron magnetic moment, 514
electron volt, 513
elementary charge, 513
elementary events, 468
elementary surface, 226
energy. *See* internal energy
energy scheme, 28–29
enhancement factor, 421
ensembles: canonical, 81–83; generalized, 85–88, 93; grand canonical, 93–95, 100; ideal gas in canonical, 84–85; microcanonical, 76; new, 86; *p-T* (pressure-temperature), 89–91; quantum, 92–93; variational derivation of, 97–99
enthalpy, as thermodynamic potential, 36, 283, 295–97, 311, 383
entropy, 15–16; axiomatic foundation of principle of, 24–26; canonical ensemble and, 81–83; definition of, 68; expressions of, 63–64; generalized ensembles and, 85–88; Gibbs formula for, 95–97; of ideal gas, 73–75; and maximum entropy (MaxEnt) principle, 99; of mixing, 295; Onsager reciprocity relations and, 376–78; as phase-space volume, 62–64; Shannon, 390–91, 504; stochastic, 390–93; systems in contact, 71–72; and variational principle, 72–73; variation of, 23
entropy postulate, 15; application of, 16–19
entropy production: adiabatic and nonadiabatic, 402–4; nonequilibrium steady states and, 380–82; two-state system and, 409–10; uncertainty relations and, 412–14
entropy production rate: average, 395–96; Schnakenberg formula for, 411
equal-area construction, 159
equations of state, 5–6, 26, 44–45; compressibility, 269; for fermion gas, 122; for ideal gas, 75, 91, 95, 99–100, 122, 281; for Lennard-Jones fluid, 265; virial expansion of, 281, 288
equilibrium, Markov process, 388
equilibrium distribution, 322
equilibrium states of thermodynamic systems, 14
equipartition theorem, 103–4; in molecular dynamics simulations, 309
ergodic matrix, 321, 509
escape rate, 359, 387
Euler equation, 41, 56–57, 93, 100
Euler gamma function, 74, 108, 284, 452, 501–2
Euler integrals, saddle-point method and, 499–500
Euler-MacLaurin series, 151
events: definition of, 468; independent, 471; mutually exclusive, 468
evolute, as phase-space region, 65
exact differential, 489
exchange integral, 165
excluded volume interaction, 418. *See also* hard core
extensive variable, 13–14; entropy and, 15–16, 17, 23, 24–26; Koenig-Born diagrams and, 37–38; observables and, 61–62; quantum mechanical representation and, 69; work and, 20
extensivity, 15
extent of reaction, 362

Faraday constant, 513
fast variables, 442
feedback protocol, 405
Fermi distribution, variational derivation of, 128–29
Fermi energy, 123–24, 128; of electron, 136
Fermi factor, 121, 124
fermion(s), 2, 118; in two dimensions, 128
fermion gas, 2, 117–22; and electrons in metal, 123–27; equation of state for, 121–22; pressure-energy relation for, 127–28
fermion system, 144
Fermi statistics, 139–40; variational derivation of, 139–40
Fermi temperature, 124
ferromagnets: critical exponents of, 205; Curie point in, 55; model of ferromagnetism and, 3. *See also* antiferromagnet; Ising model
Fibonacci sequence, 178
Fick's equation, 348
fields, 30
fine structure constant, 513
finite lattice method, 229–31
finite-size scaling, 329, 331
first Brillouin zone, 123, 202, 233–34, 239, 497–98
first-order transitions, 51
first principle of thermodynamics, 20
Fisher exponent, 431

fixed point(s), 220; cubic, 252–53; of decimation, 220; fixed lattice method and, 231; Gaussian, 40, 236, 242, 247; Heisenberg, 243, 246, 247, 249, 253; infinite temperature and, 236; Kadanoff transformation and, 225–29, 254; ϕ^4 model and, 240
Flory argument, linear polymers in solution and, 421–22
fluctuation(s): Einstein theory of, 196–99; free energy of, 197; internal energy of, 197, 198; Kadanoff transformation and, 215, 218; mean-field theory and, 199–200; of number of particles in grand canonical ensemble, 93–95; of small macroscopic systems, 196–99; of uncorrelated particles, 99–101
fluctuation-dissipation relation, 88; in canonical ensemble, 84; in generalized ensemble, 85; for magnetization, 88; Monte Carlo simulation and, 333; in *p-T* ensemble, 89–90
fluctuation-dissipation theorem, 48, 372–76; dynamics, 372–76; equilibrium, 386; relation between fluctuation and response, 374
fluctuation relations, 397–402; Crooks as, 400; detailed, 398; Gallavotti-Cohen as, 399; Hamiltonian system as, 401; integral, 397; involution property and, 398; Jarzynski equality and, 399, 402; Onsager reciprocal relations as, 400–401; and work as sufficient statistics, 401–2
fluids, classical, 262; colloidal suspensions and, 301–4; complex, 262; density functional approach and, 276–80; and fluid-fluid interfaces, 279–80; Lennard-Jones potential and, 264–65; and liquid solutions, 295–300; partition function of, 262–65. *See also* reduced densities
fluxes, 378; affinities and, 378–80
Fokker-Planck equation, 359, 367, 369; diffusion processes and, 352–53
formal neurons, 462–66
Fortuin-Kasteleyn representation, 436
forward Kolmogorov equation, 353
forward protocol, 396
forward trajectory, 396
Fourier series, 492–94
Fourier transform, 494–96
fractal, 346
fractal dimension, 347; Brownian trajectory and, 347; concept of, 347n2
fractal properties: of Brownian trajectories, 345–47; and structure of percolation cluster, 439–41
free energy: of fluctuation, 197; Landau theory of, 192–93; Legendre transformation and, 277; Maxwell relations and, 33–35; of percolation, 48; per site for one-dimensional Ising model, 219–20. *See also* Gibbs free energy; Helmholtz free energy; Landau free energy
free energy difference, 197
frozen disorder, 416
frustrated spins, 148
frustration, 417, 447–48; Hopfield model and, 464
fugacity, 120; definition of, 95, 168; as function of density, 288; grand canonical ensemble and, 420; grand canonical partition function and, 178, 264, 284–86; linear polymers and, 421; in partition function, 127, 165, 259–60; and polydisperse solution, 425; virial expansion and, 288
functionals, 98
functional derivatives, 489–91
functions defined on lattice, 496–98
fundamental equation of thermodynamics, 26–27
fundamental hypothesis of thermodynamics, 14
fundamental physical constants, 513–14
fundamental postulate, 59; and entropy as phase-space volume, 62–64

Gallavotti-Cohen relation, 399; and symmetry, 400–401
gas constant, 6, 513
gases: chemical equilibria in, 154; in gravity field, 7; internal degrees of freedom of, 146–51; van der Waals equation for, 157–61; virial expansion and, 265. *See also* boson gas; fermion gas; hard sphere gas; ideal gas; simple fluids
gas-liquid system, critical exponents of, 205
Gastheorie (Boltzmann), 262
Gaussian distribution: Brownian motion and, 340; critical exponents and, 240; and energy levels in random-energy models, 452–53; of random force, 359, 371; of random variables, 449, 452, 469; of simple fluid model, 313; of velocity, 5
Gaussian model, 419; critical crossover in, 250–51; fixed points in, 252; linear polymers in solution and, 418–21; renormalization and, 234–37
Gaussian random variable, 476
generalized Brownian motion, 365–68
generalized detailed balance, 390
generalized ensembles, 85–88; and average values and fluctuations, 87; entropy in, 95–97; grand canonical ensemble as, 93–95 (*see also* grand canonical ensemble); and paramagnet, 87–88; variational principle for, 97–99.
generalized forces, 20, 26, 199; as affinities, 378–80
Gibbs condition, 55–56
Gibbs correction factor, 74, 263
Gibbs-Duhem equation, 42, 51, 57–58, 159–60, 288, 296, 383
Gibbs formula, for entropy, 95–97
Gibbs free energy, 35–36, 41, 119; chemical potential and, 49; equilibrium and, 363; isotherm of, 50, 51; liquid solutions and, 295; partition function and, 89; simple fluid and, 44; variation of, 48

Gibbs functional S, of probability distribution, 97
Gibbs isotherm, 57
Gibbs paradox, 75
Gibbs phase rule, 54
Ginzburg criterion, 196, 247, 253; fluctuations and, 199–200
glasses, 13, 269. *See also* spin glasses
grand canonical ensemble, 93–95, 100, 163–67, 178, 420; for adsorption, 144; for boson or fermion gas, 118–19; fluctuations of number of particles in, 93–95; for simple fluid, 264, 284
gravitational constant, 514
gravitational energy, 137
gravity field, gas in, 7
Green-Kubo formulas, 344
Gumbel distribution, 450–52

Hamiltonian, 68, 69, 74, 99; critical surface and, 226; density of, 234–35; and Ising model, 164–65; as operator, 92
Hamilton's equation of motion, 67, 401
hard core, 91, 163, 264, 267, 283, 292–94
hard sphere gas: equation of state for, 288–89; one-dimensional, 90–91; virial coefficients for, 292, 293
harmonic oscillators: Brownian motion of, 375; classical, 102–3; equipartition theorem and, 103–4; Planck radiation law and, 106–9; quantum mechanical, 70, 104–5
harmonic potential, Einstein condensation, 143
Harris criterion, 444
Hatano-Sasa relation, 404
Hausdorff dimension, 439–41
heat: definition of, 20; enthalpy and, 36; and heat flux, 21; infinitesimal, 20; solvation, 296–99; stochastic, 389–92, 395; work and, 19–24
heat reservoir. *See* reservoir, thermal
Heaviside's theta function, 61; Fourier transform of, 373
Heisenberg fixed point, 243, 246, 247, 253, 444
Heisenberg model: critical exponents and, 243; Ising model and, 171, 445; and renormalization, 231
helium-4, λ transition of, 3, 55, 142, 205
Helmholtz free energy, 93, 119, 237, 300; approximations to, 277–79; availability and, 197, 366; definition of, 30; equation for, 56; fluctuation relations and, 399; partition function and, 147, 213; perturbation theory of fluids and, 294; phase coexistence and, 50, 51; in thermodynamics, 33, 277; virial expansion and, 282
heme, 145
hemoglobin, 145; simplified model of, 145–46
Hermitian operator, 69, 92
hexatic phase, 260
Hilbert space, 69, 92, 104
homogeneous and isotropic, 11
Hooke's law, 58, 89, 102–3
Hopfield model: of associative memory, 462–66; capacity of, as network, 466
hydrogen, ortho- and para-, 152–53
hypernetted chain equation (HNC), 291–92, 293

ideal gas, 2; in canonical ensemble, 84–85; classical, 144; entropy of, 73–75; and equation of state of Bose gas, 142; and equations of state, 75, 91, 95, 99–100, 122, 157, 281; free energy of, 282; in grand canonical ensemble, 94–95; molecular dynamics and, 311; in one dimension, 90–91; pressure and internal energy density in, 127; as simple fluid, 45; specific heats of, 85, 103; velocity distribution in, 77
impurities, systems with, 442–45
independent events, 471, 472
independent random variables, 448; central limit theorem for, 474; generating function of, 424
independent thermodynamic variables, 153
indeterminacy principle, 70
infinite temperature fixed point, 236
information: copying, 410–12; and information reservoirs, 407–10; thermodynamics of, 404–7
instantaneous stationary distribution, 402
integrable function, 494
integral fluctuation relation, 397, 399
integrating factor, 489
intensive variables, 30–33; fluctuations of, 197; Gibbs-Duhem equation and, 41–42; Koenig-Born diagram and, 37–38; phase coexistence and, 49–50; thermodynamic potentials and, 30–33, 38–39
interaction energy among macroscopic systems, 13–14
interfaces: fluid-fluid, 279–80; planar, 55–57
interface width, 55
internal degrees of freedom: atomic gases and, 146–48; molecular gases and, 148–51; ortho-hydrogen and, 152–53; para-hydrogen and, 152–53
internal energy: of black body radiation, 110; entropy and, 15–17; fluctuations of, 83–84, 197, 198; of ideal gas, 73; magnetization and, 197; relation between pressure and, 127–29; specific heat and, 103; variance of, 83; volume and, 89; and union of two systems, 13–14; work and, 20
internal partition functions, 147
inverse Fourier transform, 494
involution property, 398
ionic solutions, 299–300
irreducible channel, 290
irreducible diagrams, 287
irrelevant perturbations (operators), 225–29
Ising antiferromagnet: mean-field theory of, 188; order parameter and, 171

Ising model, 3; and analogy between magnetic phenomena and percolation, 427–28; critical exponents and, 206; *d*-dimensional cubic lattice and, 322–23, 324; decimation in the one-dimensional, 218–20, 221; decimation in the two-dimensional, 222; duality in the three-dimensional, 182; with exchange integral, 165; fixed point and, 252; fluctuations in, 196–97; Heisenberg model and, 171; magnetization and, 169; and Mattis model as infinite-range, 448; mean-field theory and, 182–86, 188; Migdal-Kadanoff renormalization of, 439; Monte Carlo simulations of, 333–36; neural network and, 464; one-dimensional, 176–79; of paramagnet, 78–80; partition function of the two-dimensional, 209–14; phase transitions and, 164–65; Potts generalization of, 434–36; spontaneous magnetization and, 185; square lattice with L x L spin and, 173, 174; two-dimensional, 179–82, 220–24; XY model and, 171. *See also* critical behavior of Ising model; ferromagnets
isotherm, 50–51; critical, 245; in van der Waals model, 156, 160; at zero temperature, 68, 128
isothermal compressibility, 90; at critical point, 273; pair distribution and, 269
isothermal magnetic susceptibility, 88
isotropy, 267, 340, 357

Jarzynski equality, 399; close to thermodynamic limit, 402
Jensen's inequality, 186, 481–82
joint probability, 471; and distribution, 471; Gaussian, 471
Josephson constant, 513
Joule-Thomson process, 283–84
jump rate, *394*, 387

Kadanoff transformation, 225–28, 234; and critical exponents at first order in ε, 241; finite lattice method and, 229–31; in Fourier space, 231–45; Gaussian model and, 234–37; and one-dimensional Ising model, 218–20; in ϕ^4 model, 238–39; and relevant operators, 225–28
Kappler's experiment, 103, 353
kinetic coefficients, matrix of, 380–82
kinetic energy, as observable, 61; temperature and, 309
Kirkwood equation, 275
Klein-Kramers equation, 353
Knudsen gas, 7
Koch curve, 440, 441
Koenig-Born diagram, 37–38
Kosterlitz-Thouless transition, 255–61
Kramers-Kronig relation, 374
Kronecker delta, 163, 237, 433, 456; definition of, 116
Kubo formula, 358; sum rules and, 357–58
Kullback-Leibler divergence, 391, 506

Lagrange multipliers, 98, 99, 484–86
λ transition, 3, 55, 142, 205
Landau argument, one-dimensional Ising model and, 176
Landau diamagnetism, 131–33
Landauer's principle, 404–5
Landau free energy, 193, *193*
Landau levels, 132
Landau theory of phase transitions, 192–95
Langevin diamagnetism, 129–31
Langevin equation, 341, 367; linearized, 368
Langevin function, 133
Langevin theory of paramagnetism, 88
Langmuir isotherm, 144, 145
Larmor theorem, 131
Lattice gas, 163; phase coexistence and, 162–64
law of adiabatic processes, 47
law of corresponding states, 159
law of large numbers, 475
law of mass action, 362
law of total probabilities, 468
Lee-Yang theory of phase transitions, 165–69; electrostatic analog of equation of state and, 168–69; and phase coexistence, 167–68; theorems of, 166
Legendre transform, 482–84; enthalpy and, 36; entropy and, 86; Lagrange multipliers and, 484–86; number of particles and, 38; properties of, 50; thermodynamic potential and, 30, 32
Lennard-Jones potential, 264–65; gas expansion and, 283, 284; Mayer function of, 281–82; perturbation theory and, 294, 295
lever rule, 52, 156
linear congruence (random number generator), 326
linear polymers in solution, 417–25; comparison of, to percolation problem, 432–33; critical phenomena and, 423–25; and Flory argument, 421–22; and Gaussian model, 418–21; and polydisperse solution, 425; and on lattice, 421
Liouville's equation, 67, 339
Liouville's theorem, 64–68
liquid(s): colloidal suspensions of, 301–4; entropy of, 295; perturbation theory and, 293–94
liquid-gas coexistence, 156–62, 169–70; binary mixtures of, 161–62; critical point and, 155–56
liquid solutions, 295–300; ionic, 299–300; solvation heat and, 296–99
local-density approximation (LDA), 279
logsum inequality, 482
Lorentzian correlation function, 356
lower critical dimension, 255

M (magnetic moment), 20, 42; as magnetic dipole, 11; density of, per unit of volume, 42

magnetic quantum number, 150
magnetic field, 42–43, 79, 88, 130–31, 147, 169, 177, 188, 195–97, 215–17, 226, 245, 457–58
magnetic susceptibility: critical exponents and, 196; Curie-Weiss law for, 190, 193; isothermal, 88; of spin glass, 457.
magnetization, 11; canonical ensemble and, 80; density and, 42; internal energy and, 197; Ising model and, 169; and magnetic moment (M), 43; paramagnet and, 87–88; thermodynamics and, 9, 42–44
majority rule, 230
manipulation, 386
Marcus and MacFee experiment, 8–9
Markov chain(s), 336, 507; convergence to stationarity and, 509–12; definitions of, 507–8; microcanonical ensemble and, 325; Monte Carlo method and, 321–22; non-ergodic matrices and, 512; spectral properties of, 508–9; time-homogeneous, 508
Markov process, 352; discrete, 387; equilibrium and, 388
mass action, law of, 154
Massieu functions, as thermodynamic functions, 86
master equation, 321, 387, 387n1, 392–93, 403, 408, 412, 508–9
Mattis model, 448
maximum entropy (MaxEnt) principle, 99
Maxwell-Boltzmann distribution, 311, 349, 351, 384
Maxwell construction, 159, 160
Maxwell distribution, 6, 7, 77; Marcus and MacFee experiment and, 8–9
Maxwell relations, 90, 161; free energy and, 33–35; Koenig-Born diagram and, 37–38; as thermodynamic potentials, 34
Maxwell's postulates, 4, 5
mean, of *N* independent random variables, 474
mean-field theory, 182–86; cluster approximations and, 187; correlation functions and, 188–92; critical exponents and, 206; Ginzburg criterion and, 199–200; graphical solution of, 184; Hopfield model and, 465; of Ising antiferromagnet, 188; Landau theory and, 192–95; variational principle for, 186–87
mechanical energy, 14
mechanocaloric effect, 382–83
mesoscopic systems, 386
metals, electrons in, 123–27
metastable state(s), 359; stochastic resonance and, 358–62
Metropolis algorithm, 323, 334–36, 437; one-dimensional Ising model and, 325
Meyer's relation, 47
micelles, 301–4
microcanonical ensemble, 76; canonical ensemble and, 81–82; Markov chain for sampling of, 325; probability distribution and, 75–76
microemulsions, 14, 56
microstate, 60–61, 96–100, 102, 106, 119, 139, 144, 197, 314, 323, 327, 368
Migdal-Kadanoff transformation, 438, 439
molecular dynamics, 307; numerical simulation using, 307–14; proof of canonical sampling and, 318–19; temperature and pressure in, 309–11; thermostats in, 314–19; Verlet algorithms and, 311–14; Widom insertion method and, 333. *See also* Monte Carlo method
molecular gases: internal degrees of freedom of, 148–51; as simple fluids, 262–63
moments of a random variable, 470; generating function and, 424
momentum: total, 5, 115, 315; total kinetic energy and particle, 123
momentum density, 113
monomers, 417
monotonicity of entropy, 16
Monte Carlo method, 307; accelerated, 333–36; and algorithms in statistical mechanics, 322–25; dynamics of, 464; extrapolation to thermodynamic limit and, 329–31; Markov chains and, 321–22; Metropolis algorithm and, 323; microcanonical ensemble and, 325; numerical simulation and, 319–31; random sequences and, 325–27; statistical errors and, 327–29; Widom insertion method and, 333
mutual information, 405, 506
mutually exclusive events, 468
myoglobin, 145

N-dimensional spaces, properties of, 502–3
nearest neighbors, number of, 165. *See also* coordination number
Nernst's postulate, 68, 69, 70n1
neural networks, 2, 464
neuron, 462, 463
neurotransmitters, 462
neutrino density, in universe, 128–29
neutron scattering experiment, 270–71
neutron stars, 138–39
Newton's law, 113, 114
node of diagram, 286
nonadiabatic entropy production, 402
noncorrelation, 340
nonequilibrium entropy, 391
nonequilibrium states, 1
nonequilibrium steady states, 380–82
non-ergodic matrices, 512
nonobservable, 61
normalization condition, 476

normalization constant, 86, 87
normalization factor, 92
Nosé-Hoover thermostat, 315
nuclear degrees of freedom, 146–48
nuclear magneton, 514–18
numerical experiments, 306–7
numerical simulation: accelerated Monte Carlo methods and, 333–36; discussion of, 336–37; introduction to, 305–7; molecular dynamics as method of, 307–14; Monte Carlo method of, 319–31; and umbrella sampling, 331–33

objective (ontic) probabilities, 97
observables, 61–62, 72; generalized Brownian motion and, 365; molecular dynamics simulation and, 337; quasi-thermodynamic, 265
occupation number(s), 99, 106, 118
one-dimensional Ising model, 176–79; decimation of, 218–20; and Landau argument, 176; solution of, 176–77
Onsager's reciprocity relations, 372, 376–78, 400–401; application of, 382–84; equilibrium and, 386; and mechanocaloric effect, 382–83; and thermomechanic effect, 383–84
Onsager's regression hypothesis, 372
order parameter, 170–72; Edwards-Anderson, 446; equation of state for, 205; and Landau theory, 194–95; in limit $n \to \infty$, 201
Ornstein-Zernike theory, 188; approximation using, 199, 421; Debye-Huckel theory compared to, 300; Gaussian model and, 236
ortho-hydrogen, degrees of freedom and, 152–53
osmotic pressure, 24, 39–40, 296
overlap between spin configurations, 328, 454, 458–59, 465–66

packing ratio, 289, 292–93
pair density, 266
pair distribution *g(r)*, 266–69; integral equations for, 289–93; isothermal compressibility and, 273; of Lennard-Jones fluid, 294, 295; Mayer function and, 281; measurement of, 270–72; reversible work theorem and, 269–70; superposition approximation using, 274
pair potential, 157, 163; in integral equations, 289–93; in numerical simulation, 307; for simple fluid, 262, 264
para-hydrogen, degrees of freedom, 152–53
parallel connection diagram, 290
parallel tempering, 336
paramagnet: generalized ensembles and, 87–88; Ising, 78–80
paramagnetism: Langevin theory of, 88; Pauli, 126–27
parity of observable, 369
partial derivatives, 487–89
particles: phonons as, 113; with variable spin, 147–48
partition function, 82–84, 104; of classical fluids, 262–65; grand canonical, 118–19, 127; hard sphere gas and, 91; integral defining, 83; in *p-T* ensemble, 89; single-particle, 119; thermodynamic potentials and, 83, 92, 96; translational, 147; of two-dimensional Ising model, 209–14
patterns (memory), 465
Pauli exclusion principle, 3, 118, 148, 152, 264
Pauli paramagnetism, 126–27
Peierls's argument, 172–76, 178
pendulum, Verlet algorithm for, 312–13
percolation, 416, 425–41; on Bethe lattice, 428–32; and fractal structure of percolation cluster, 439–41; limit $q \to 0$ and, 436–37; limit $q \to 1$ and, 434–36; and magnet phenomena analogy, 427–28; Migdal-Kadanoff renormalization of Ising model, 439; in one direction, 28; Potts model and, 432–37; renormalization of bonds in, 437–39; site, in two dimensions, 426; and standard problem, 426; theory of, 426; tree statistics and, 436–37
percolation cluster: fractal structure of, 439–41; incipient, 441
percolation threshold, 417, 426; for Bethe lattice, 428–32; fractal dimension at, 439–41
Percus-Yevick equation, 292, 293
perturbation theory, 293–95
phase coexistence, 49–51, 167–68; Clausius-Clapeyron equation and, 51–52; and coexistence curve, 52–53; critical point and, 54–55, 155–56; and Gibbs phase rule, 54; liquid-gas and, 155–56; with planar interfaces, 55–57; of several phases, 53–54; and thermodynamics, 49, 51–52. *See also* coexistence curve; critical behavior
phase diagram, 54, 156–57, 185, 195, 262, 409
phase space, 9, 59–61, 71, 81–82; accessible volume of, 74; entropy and, 62–64; Liouville's theorem and, 64–68; probability distribution and, 75–76, 86; unit of volume in, 69–71
phase transitions, 51–52; binary mixtures and, 161–62; continuous, 51, 53, 54; discontinuous, 51; duality and, 179–82; Ising model and, 164–65; lattice gas and, 162–64; Lee-Yang theory of, 165–69; and liquid-gas coexistence and critical point, 155–56; mean-field theory and, 182–86; one-dimensional Ising model and, 176–79; order parameter and, 170–72; Peierls argument and, 172–76; symmetry breaking and, 169–70; universality and scaling and, 204–9; van der Waals equation and, 157–61
ϕ^4 model, 237–40; cubic anisotropy and, 252–53
phonons, 111; second sound and, 113–17; specific heat of solids and, 111–12

photons, 3, 108; of black body radiation, 106–11
physical constants, 513–14
planar interfaces, 55–57
Planck constant, 513
Planck's law, of black body radiation, 106–9, 110
plaquettes: dual lattice of, 180–82; frustrated, 447
Poisson-Boltzmann theory, 299
Poisson bracket, 67, 68
Poisson distribution, 469
polydisperse, 301, 425
polylog function, 120–21, 141
polymers, 410–11, 417, 423–25; heterogeneous, 417; homogeneous, 417; on a lattice, 421
potential well, phase space volume for, 71
Potts model, percolation and, 432–37
power, and fluctuation-dissipation theorem, 473
pressure, 18, 22–23; Carnahan-Starling extrapolation formula and, 293; enthalpy and, 35–36; in molecular dynamics, 309–11; osmotic, 24; pair distribution and, 268; relation between internal energy and, 127–29
pressure equation, 268
principal value, 373
principle of minimal entropy production, 382
principles of thermodynamics, 11
probability: conditional, 471; definition of, 468; law of total, 468
probability density function, 469
probability distribution, 469; Gibbs functional, 97–98; of microcanonical ensemble, 75–76; phase space and, 86
product rule, 491
products, of chemical reaction, 48
protocol, 389
proton magnetic moment, 514
proton mass, 514
p-T (pressure-temperature) ensemble, 89–91, 92; one-dimensional hard sphere gas and, 90–91
pure states, 172, 417

quadratic anisotropy, 246–47
quantum ensembles, 92–93
quantum exclusion principle, 152
quantum gases, 2
quantum harmonic oscillators, 104–5
quantum mechanics, particles, 118
quantum number: angular, 150; magnetic, 150
quantum state: particle in potential well and, 71; in thermodynamics, 68–71
quasi-average, magnetization and, 170
quasi-long-range order, 256
quenched average, 443
quenched disorder, 416
quenched variables, 442–43

radiation, black body, 3, 106–11
radius of gyration, 419
random-energy model, 448–52; as limit of spin model, 452–54; replica method approach to, 454–57
random field, 445
random sequences, Monte Carlo method and, 325–27
random variables: average, 470, 471; continuous, 469; definition of, 469, 474; discrete, 469; independent, 471–73; moments of, 470–71; probability distribution and, 471; uncorrelated, 472
rate constants, 363
rate equations, 363
Rayleigh chain, 418
reactants, 48, 154
reaction coordinate, 364
realization of disorder, 443
reciprocal lattice, 497
reduced critical temperature, for Lennard-Jones fluid, 265
reduced densities, 265; BBGKY hierarchy and, 273–75; direct correlation function and, 273; for Lennard-Jones fluid, 265–66; measurement of *g(r)* and, 270–72; pair distribution and, 266–69; reversible work theorem and, 269–70
reduced variables, 158
Reech's relation, 47
refractory period, 463
relative entropy, 391, 506
relevant perturbations (operators), 225–29
renormalization: 1/n expansion of critical exponents and, 245; bond percolation and, 437–39; critical exponents at first order in ε and, 240–44; in Fourier space, 231–45; Gaussian model and, 234–37; Kosterlitz-Thouless transition and, 255–61; Migdal-Kadanoff, of Ising model, 439; multiplicity of contributions to, of u, 244–45; ϕ^4 model and, 237–40; two-dimensional Ising model and, 224. *See also* critical exponents; Kadanoff transformation
renormalization group, 215
replica method, 417, 444; and random-energy model, 454–57; Sherrington-Kirkpatrick model and, 457–62
replica symmetry breaking (RSB), 455–57; Parisi scheme and, 461–62
representative point, 59
reservoir, thermal, 85; canonical ensemble and, 81; and heat, 24; information reservoir as, 407–10
response functions, 368–72; equation defining, 370
reversible adiabatic process, 24
reversible work theorem, 269–70
Riemann zeta function, 108

root: of Bethe lattice, 428–32; of diagram for evaluating virial coefficients, 286, 287
rotational degrees of freedom, 149, 150
runaway, 253
running average, 328

Sackur-Tetrode formula, 75
saddle-point method, 192, 202–3, 499; equations for, 202–3
Sagawa-Ueda relation, 406, 409
sample space, 468
sampling with repetition, 140
scale invariance, 25
scaling, universality and, 204–9
scaling laws, 207–9; finite-size scaling and, 329, 331; Kadanoff transformation and, 225; linear polymer and, 421
Schnakenberg formula, 396; entropy production rate from, 411
Schrödinger equation, 69, 132; for nuclei, 149
second-order transitions. *See* phase transitions: continuous
second principle of thermodynamics, 16, 25, 379
second sound: phonons and, 113–17
second sound speed, 117
semiconductors, doped, 416
semifactorial, 501
semipermeable membrane, 24
semipermeable walls, 14
series diagrams, 290–91
Shannon entropy, 97, 99, 390–91, 504; information reservoirs and, 407; of a probability distribution, 96
Sherrington-Kirkpatrick model, of spin glass, 457–62
Sierpiński triangle, 439–41
simple fluids, 2, 11, 41; chemical potential for, 39–40; equations of state for, 44–45; Gibbs free energy for, 35, 37, 49; numerical simulation of, 313–14; partition function for, 262–65; phase coexistence in, 49–51; phase space for, 59–61
simulated annealing, 334; protocol of, 334–35
simulation, 305; as third branch of science, 306. *See also* numerical simulation
sine series, 493
single-particle potential, 74, 84, 263, 277–78
singularity, 257; strength of a, 257
site percolation, 426
Smoluchowski equation, 351; Brownian particle and, 347–52
soft matter, 262
solubility, 297
solute, 23–24, 39–40, 57, 295–99
solution, linear polymers in, 417–25
solvation heat, 297; liquid solutions and, 296–99
solvent, 23–24, 39–40, 57, 295–99, 301, 339–43, 352–54; good, 418
spanning trees, 436
species (chemical), 362
specific heat: critical exponents and, 207; critical temperature and, 156; energy fluctuations and, 83–84; of ideal gas, 85, 103; of intensive vs. extensive variable, 13–14, 30–33; in Ising model, 213, 333; of Ising paramagnet, 78–80; in Landau theory, 194; numerical simulation and, 309; rotational degrees of freedom and, 149, 150; of solids and phonons, 111–12
speed of light, 513
speed of sound, 114
spherical model, 204
spin, 3; frustrated, 148; particles with variable, 147–48; statistics and, 118
spin glasses, 417, 445; frustration and, 447–48; Hopfield model and, 462; random-energy model and, 448–52; random-energy model as limit of, 452–54; replica method approach to random-energy model and, 454–57; replica symmetry breaking (RSB) and, 455, 457; Sherrington-Kirkpatrick model and, 457–62
spontaneous magnetization, 170, 185; Ising model and, 185
spontaneous symmetry breaking, 169; order parameter and, 170–72; Peierls argument and, 172–76
square gradient approximation, 279
stability: adiabatic process and, 25; thermodynamic equilibrium and, 45–47
staggered magnetization, 171
standard deviation, 445, 471
standard gravity acceleration, 514
standard molar volume, 513
stars: neutron, 138–39; white dwarfs, 135–39
states of equilibrium. *See* thermodynamic equilibrium
stationary distribution, 322, 388, 509
stationary Markov chains, 509–12
statistical errors: Monte Carlo method and, 327–29; and mutual overlap, 328; and running average, 328
statistical mechanics, 1–3; Monte Carlo algorithms in, 322–25
statistical postulates, 3–4
statistics: of loops, 210–12, 214; variational derivation of Bose, 140–41; variational derivation of Fermi, 139–40
steady-state distribution, 388
steady states, nonequilibrium, 380–82
Stefan-Boltzmann law, 110
Stevin's law, 47
Stirling formula, 74, 78, 84, 139, 501

stochastic differential equation, 341
stochastic energetics, 390; thermodynamic consistency and, 389–90
stochastic matrix, 321, 508
stochastic mutual information, 406
stochastic resonance: metastable states and, 358–62; phenomenon of, 361–62
stochastic thermodynamics, 386; adiabatic and nonadiabatic entropy production and, 402–4; average entropy production rate and, 395–96; copying information and, 410–12; fundamentals of, 386–89; information and, 404–7; information reservoirs and, 407–10; simple examples of, 393–95; time reversal and, 396–97; uncertainty relations and, 412–14
stochastic work, 389, 395
stoichiometric coefficients, 48, 49
structure factor, 272, 278, 292, 293
subjective probabilities, 97
sum rules, 358; Kubo formula and, 357–58
superposition approximation, 274
surface tension, 56, 57, 280, 301
surfactants, 57, 301
symmetry, Hamiltonian, 165, 170
symmetry breaking, 169, 173; phase coexistence and, 169–70. *See also* spontaneous symmetry breaking
synapses, 462; exciter, 463; inhibitor, 463
systems with impurities, 442–45

temperature: absolute, 20–22, 88; as conjugate of entropy, 33–34; in molecular dynamics, 309–11
tensioactives, 57. *See also* surfactants
theory of elasticity, 57–58
thermal, variables, 442
thermal de Broglie wavelength. *See* de Broglie thermal wavelength
thermal equilibrium, 21
thermally conductive (diathermal) walls, 14
thermodynamic consistency, 390; and compatibility, 386; stochastic energetics and, 389–90
thermodynamic equilibrium, 1, 12; chemical reactions and, 48–49; stability and, 45–47; states of 11–13; of systems in contact, 71–72
thermodynamic experiment, generic diagram of, 15
thermodynamic limit, 62; Monte Carlo method and, 329–31
thermodynamic potential(s), 30–33; equation for, 39; Koenig-Born diagram and, 37–38; Legendre transform and, 30, 35, 36, 38; stability of thermodynamic equilibrium and, 46; of systems in contact, 71–72
thermodynamic properties, 12–13
thermodynamics: central problem of, 14–15; first principle of, 20; fundamental hypothesis of, 14; of information, 404–7; magnetic systems and, 42–44; Onsager's relations and, 382; quantum states and, 68–71; second principle of, 16, 25; stochastic fundamentals of, 386–89
thermodynamic systems, 11–13
thermodynamic uncertainty relation, 414
thermomechanical effect, 383–84
thermostats, molecular dynamics, 314–19
theta function, 61
Θ-point, 418
third law of thermodynamics, 68
three dimensional $n \to \infty$ model, critical exponents for, 206
tie lines, 53
time reversal, stochastic thermodynamics and, 396–97
timescale separation, 386
time-translation invariance, 343, 355, 356, 370
trajectory x, 387–88
transfer matrix, 177
transition matrix, 321
transition pressure, 156
transition probabilities: of Markov chain, 321, 507; in Monte Carlo method, 319
transition rate, 387
transition temperature, for Einstein condensation, 142
transitivity, of adiabatic equivalence, 25
trees on a lattice, 436–37
tricritical point, 195
triple point, 53, 156, 157, 295; of water, 53
Trotter's formula, 317
two-dimensional Ising model: decimation results in, 220–24; partition function of, 209–14; symmetry breaking in, 172–76
two-tape system, 409–10

ultrametricity, 462
umbrella sampling, 331–33
uncertainty relations, 412–14
uncorrelated particles, fluctuations of, 99–101
uncorrelated variables, 472
uniform distribution, 469; in phase space, 76
universality and scaling, 204–9, 226, 231
universality classes, 195, 205
upper critical dimension, 254

vacuum permeability, 513
vacuum permittivity, 513
van der Waals equation, 157–61
van Koch curve, 440, 441
van 't Hoff's law, 296, 299; and thermodynamics of diluted solutions, 349
vapor pressure, 51, 156, 159
variance, 54, 471, 472; of internal energy, 83; of magnetization, 88; of number of particles, 100

variational derivation, of ensembles, 97–99
variational principle, 484; entropy and, 72–73; of Legendre transform, 277; mean-field theory from, 186–87
vector order parameter, 171; critical exponents and, 245; Landau theory and, 194–95; in limit $n \to \infty$, 201–4; for renormalization in Fourier space, 244–45
Verlet algorithms, 311–14
vibrational degrees of freedom, 149, 150
virial coefficients, 281; first, 281; hard-sphere gas and, 292, 293; second, 283, 284; subsequent, 284–89
virial expansion, 265, 281–93; of equation of state, 288; of integral equations for the *g(r)*, 289–93
virial theorem, 310
virtual states, 15
viscosity: Brownian motion and, 340, 341, 343; constitutive relations and, 354
volume: of phase space, 69–71, 263
von Klitzing constant, 513
vortex, 256, 257

walls, 14–15; adiabatic, 14, 23; mobile, 18–19, 23; semipermeable, 14, 19, 23–24, 39, 296; thermally conductive, 14, 16, 18–19
wavenumber, 494
wave vector, 143
weight, 232
white dwarfs, 135–39
Widom insertion method, Monte Carlo simulation and, 333, 402
Wien's law, 109
work, 19–24: fluctuation-dissipation theorem and, 374; heat and, 19–24; Helmholtz free energy and, 30; mechanical, 20, 23; and reversible work theorem, 269–70; sign convention for, 20; stochastic, 385, 389–92

XY model: critical exponents for, 206; definition of, 171; in two dimensions, 255–61

zero-mass particles, Einstein condensation of, 143
zero-point energy, 105
zeta function, 108